

Review

Design and application of lanthanide chelating tags for pseudocontact shift NMR spectroscopy on biomacromolecules

Daniel Joss, Daniel Häussinger*

*corresponding author

Affiliations:

Daniel Joss (daniel.joss@unibas.ch), PD Dr. Daniel Häussinger (daniel.haeussinger@unibas.ch, +41 61 207 11 59)

University of Basel

St. Johannis-Ring 19

4056 Basel

Switzerland

Number of pages: 60

Number of figures: 29

Number of tables: 1

Keywords: paramagnetic NMR, pseudocontact shift, residual dipolar coupling, lanthanide chelating tag, biomacromolecule

Declarations of interest: none

Abstract

In this review, lanthanide chelating tags and their application in pseudocontact shift NMR spectroscopy as well as analysis of residual dipolar couplings are covered. Including a complete overview of non-DOTA-derived and DOTA-derived lanthanide chelating tags, critical points in the design of lanthanide chelating tags as appropriate linker moieties, stability under reductive conditions e.g. for in-cell applications, magnitude of the transferred anisotropy from the lanthanide chelating tag to the biomacromolecule under investigation and structural properties as well as conformational bias of the lanthanide chelating tags are discussed. Furthermore, all currently published DOTA-derived lanthanide chelating tags used for PCS NMR spectroscopy are displayed in tabular form including their anisotropy parameters with all employed lanthanide ions, C_B-Ln distances and tagging reaction conditions, i.e. stoichiometry of lanthanide chelating tag, pH, buffer composition, temperature and reaction time. Additionally, reported applications of lanthanide chelating tags for pseudocontact shifts and residual dipolar couplings on proteins, protein-protein and protein-ligand complexes, carbohydrates, carbohydrate-protein complexes, nucleic acids and nucleic acid-protein complexes are presented and critically reviewed. The vast and impressing field of applications of lanthanide chelating tags in structural investigations of biomacromolecules in solution clearly illustrates the significance of this particular field of research and the extension of the repertoire of lanthanide chelating tags from proteins to nucleic acids holds great promise for the determination of valuable structural parameters and further developments in characterising intermolecular interactions.

1. Table of contents

2.	Introduction.....	3
2.1	Paramagnetic nuclear magnetic resonance spectroscopy on biomacromolecules.....	3
2.2	Pseudocontact shift	6
2.3	Residual dipolar couplings	8
3.	Lanthanide chelating tags.....	10
3.1	Introduction & Overview	10
3.2	Non-DOTA-based lanthanide chelating tags.....	11
3.3	DOTA-based lanthanide chelating tags	16
3.4	Tabular overview of published DOTA-derived LCTs	25
3.5	Parameters in the design of LCTs	32
4.	Applications of lanthanide chelating tags in the structural characterization of biomacromolecules	34
4.1	Applications of lanthanide chelating tags on proteins I: pseudocontact shift	34
4.2	Applications of lanthanide chelating tags on proteins II: residual dipolar couplings	39
4.3	Applications of lanthanide chelating tags on proteins III: protein-ligand complexes	40
4.4	Applications of lanthanide chelating tags on carbohydrates	42
4.5	Applications of lanthanide chelating tags on nucleic acids	45
5.	Conclusion & Outlook.....	47
6.	Acknowledgement.....	48
7.	Glossary	48
8.	References.....	51

2. Introduction

2.1 Paramagnetic nuclear magnetic resonance spectroscopy on biomacromolecules

Nuclear magnetic resonance (NMR) spectroscopy is one of the most important analytical techniques for many different fields in chemistry and biology, from organic synthesis to structural biology. In chemistry, NMR is mainly used as a routine tool for topology confirmation of small molecules as well as 3D structure elucidation. The development of more advanced NMR techniques, i.e. homo- and heteronuclear NMR pulse sequences for establishing through-space and through-bond correlations between nuclei and suitable for complex structural studies, turned out to be crucial for the understanding of structure and function of biomacromolecules, e.g. proteins as well as deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). The distance limitation of nuclear Overhauser effect spectroscopy (NOESY) in the range of 5-8 Å and the unspecific shifts of chemical shift perturbation (CSP) experiments, however, render long-range paramagnetic effects as highly valuable, complementary tools to measure distance and orientational restraints in biomacromolecules [1-24]. Paramagnetic effects observed in NMR include paramagnetic shifts (contact (CS) and pseudocontact shifts (PCS)) [1-3, 6, 8, 10, 14, 23-31], paramagnetic relaxation enhancement (PRE) [16, 17, 22, 32-36], residual dipolar couplings (RDC) [37-45], dipolar shift anisotropy (DSA), cross-correlation between DSA and dipolar relaxation (DSA/DD) and residual anisotropic chemical shifts (RACS). The different effects associated with the introduction of a paramagnetic lanthanide ion yield complementary structural restraints and therefore present a powerful tool-box for the structure determination of biomacromolecules in solution. Whereas CSs that are through-bond effects can be neglected for studies of biomacromolecules, PCSs yield structural restraints up to distance of more than 100 Å due to their favourable distance dependence of R^{-3} and the 3D spatial information that is encoded in the induced shifts. PRE offers the opportunity to obtain information of the distance of a nucleus measured from the position of the paramagnetic centre by line broadening and hence signal intensity decrease. The distance dependence of PRE effects (R^{-6}) is less favourable than for PCSs (R^{-3}) and no angular information can be obtained, but the magnitude of the observed effect suffers less from motional averaging. RDCs provide valuable information about the relative orientation of bond vectors in space. To observe RDCs, partial alignment of the protein is required. This can be obtained by either the introduction of a paramagnetic centre or use of an alignment medium as e.g. bacteriophages, bicelles or acrylamide gels [46, 47]. Due to the favourable distance dependence of PCSs (R^{-3}) and RDCs (no distance dependence, but motional averaging can lead to observation of decreased RDCs at larger distances from the paramagnetic centre) as well as the fact, that PCSs provide not only distance but also angular information and thus yield valuable positional restraints, this review focuses on tags for the exploitation of PCSs and RDCs for the structural investigation of biomacromolecules and their complexes.

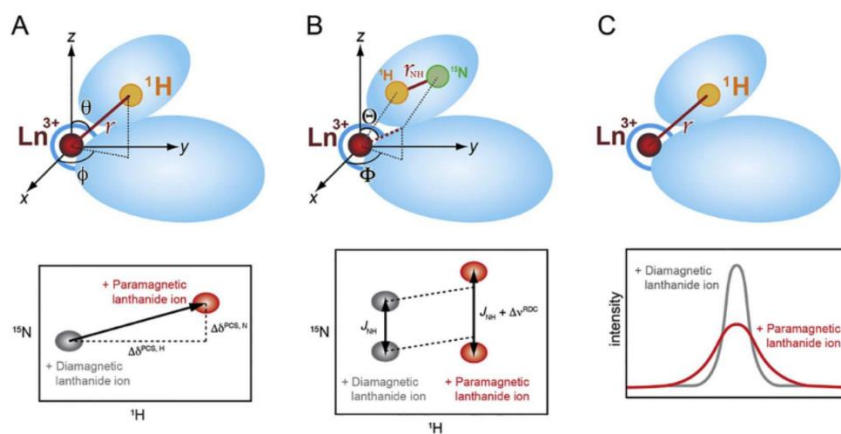


Fig. 1. Effects associated with the introduction of a lanthanide ion [48]. A) 3D spatial information obtained by PCS. B) Orientational restraints obtained by RDCs. C) Distance measurement using PRE.

Soon after a landmark study published by Barry et al. [49], who could quantitatively determine the conformations of mononucleotides in solution by scanning through molecular geometries and matching them with the lanthanide induced NMR shifts using a Ferranti Argus 500 computer, the theoretical background of the NMR shifts induced by paramagnetic lanthanides was established in the 1970s by Bleaney et al. [50, 51]. However, recently also critics and new findings arose concerning the validity of Bleaney's theory about the magnetic properties of lanthanide complexes concerning paramagnetic NMR [52, 53]. The very first applications of paramagnetic NMR include the europium shift reagents that allow to increase the shift range in order to resolve signal overlap and determine the enantiomeric ratio of chiral substances by forming a diastereomeric complex in-situ [54-58]. After the successful exploitation of paramagnetic effects of metal ions contained in proteins by Bertini et al., e.g. in zinc finger moieties [16], and determination of the associated paramagnetic susceptibility tensors, the next step in the development of paramagnetic NMR contained lanthanide binding peptides (LBPs) with micro- and nanomolar affinity to lanthanides mimicking the zinc finger motif in order to incorporate paramagnetic [15, 59] and fluorescent [60, 61] metal ions into proteins without native metal binding site. Subsequently, the development of single- and double-armed lanthanide chelating tags (LCTs) that can be attached to a protein of interest and, thus, yield valuable structural restraints for the determination of the structures of biomacromolecules in solution strongly expanded the field of paramagnetic NMR [9, 13, 14, 24, 62, 63]. Today, paramagnetic NMR is an established method in structural biology and includes many different applications, such as refining crystal structures of proteins [64], probing solution dynamics of proteins [65], localization of ligands on proteins [20, 66-71], study of protein-protein complexes [72], alleviation of signal dispersion in crowded HSQC spectra [73], facilitating analysis of large sized proteins [12], conformational analysis of carbohydrates [74], as well as investigation of the solution structure of nucleic acids [30].

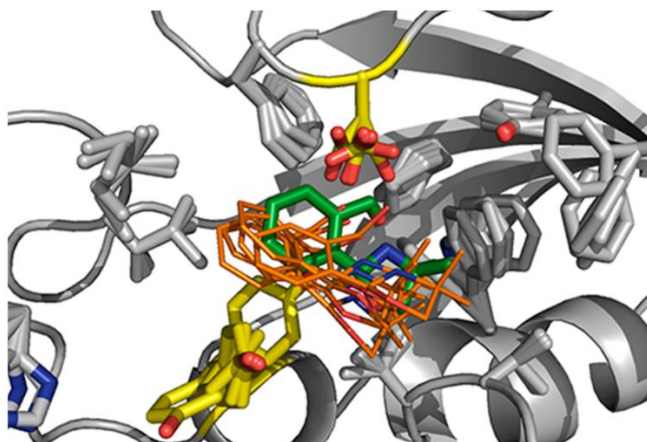


Fig. 2. Localization of a ligand on a protein by PCS (green: averaged NOESY structure, orange: best five obtained structure by PCS-fitting) [69]. (Reprinted with permission from J.-Y. Guan, P.H.J. Keizers, W.-M. Liu, F. Löhr, S.P. Skinner, E.A. Heeneman, H. Schwalbe, M. Ubbink, G. Siegal, *Small-Molecule Binding Sites on Proteins Established by Paramagnetic NMR Spectroscopy*, JACS, 135 (2013) 5859-5868. Copyright 2013 American Chemical Society.)

Besides the application of paramagnetic effects in protein NMR spectroscopy, also detailed research studies on solvent dependence as well as pH and temperature effects on the magnetic anisotropy caused by paramagnetic molecules alone and when attached to proteins were performed [28, 75].

Furthermore, computational approaches as e.g. density functional theory (DFT) calculations allow to predict the structure and conformational bias of lanthanide complexes in vacuo and in solution as well as the anisotropy tensor for electron paramagnetic resonance (EPR) studies and are therefore under current investigation in order to facilitate the use of paramagnetic effects by decreasing the experimental work associated with the development of new LCTs [23, 76-83]. Very recent work proves the semi-empirical approach to be superior in reproducing PCS compared to a quantum chemistry approach [84].

Comprehensive reviews about lanthanide chelating tags and their applications are available in the literature [1, 7, 85-91]. A very recent article by Su et al. documents his contributions to the field [92]. Furthermore, reviews were published that give an insight into strategies for measurements of PCSs in proteins [93], prospects of lanthanides in biomolecular NMR including LCTs [94], analysis of protein-ligand [20] as well as protein-protein complexes [95, 96] using PCSs and RDCs, the combination of paramagnetic NMR long-range restraints and crystallography [97, 98], sparse-labelling in combination with long-range restraints obtained by paramagnetic NMR [99] and LCTs used for the conformational analysis and interactions of oligosaccharides [100]. This review will focus on recent developments in the synthesis of LCTs and their application for PCS and RDC in structural biology, with an emphasis on DOTA-based LCTs. As the intricate fundamental physics of PCS and RDC have several times been discussed in great detail [85, 101, 102], we will just summarize the basics in the following chapters.

2.2 Pseudocontact shift

The dipolar, through-space effect on the shift of signals in NMR associated with the introduction of a paramagnetic centre is termed pseudocontact shift [101]. The shift perturbation arises from a contribution of the unpaired electrons of the paramagnetic centre to the effective magnetic field that a specific nucleus experiences. This shift perturbation is dependent on the distance and angular positioning of the nucleus to the paramagnetic centre. Considering this 3D spatial information and the distance dependence of the effect of R^{-3} , valuable structural information can be gained about the molecule in vicinity of the paramagnetic centre. The shift that a nucleus experiences can be calculated from the following equation that can be displayed in polar or cartesian coordinates [101]:

$$\text{PCS (polar coordinates)} \quad \delta^{PCS} = \frac{1}{12\pi r^3} \left[\Delta\chi_{ax}(3\cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{rh}\sin^2\theta\cos 2\phi \right] \quad \text{Eq. (1)}$$

$$\text{PCS (cartesian coordinates)} \quad \delta^{PCS}(x, y, z) = \frac{1}{12\pi r^3} \left[\Delta\chi_{ax} \frac{2z^2 - x^2 - y^2}{r^2} + \frac{3}{2}\Delta\chi_{rh} \frac{x^2 - y^2}{r^2} \right] \quad \text{Eq. (2)}$$

In order to describe the tensor of the anisotropic part of the magnetic susceptibility associated with a paramagnetic centre, e.g. of an LCT when attached to a protein, 8 parameters are used. Besides the x, y and z position of the metal centre and the three Euler angles α , β and γ that describe the orientation of the tensor with regard to the protein, the magnitude of the anisotropy of the magnetic susceptibility is described by the axial component and the shape of the tensor by the ratio of the axiality to the rhombicity of the anisotropy tensor. The axiality and rhombicity of a tensor give its directional asymmetry and are defined as follows [101]:

$$\text{Axiality} \quad \Delta\chi_{ax} = \chi_{zz} - \frac{\chi_{xx} + \chi_{yy}}{2} \quad \text{Eq. (3)}$$

$$\text{Rhombicity} \quad \Delta\chi_{rh} = \chi_{xx} - \chi_{yy} \quad \text{Eq. (4)}$$

As depicted in Fig. 3, anisotropic χ -tensors are usually displayed as isosurfaces for given PCS values. For a pronounced, high axial/rhombic ratio, a tensor resembling a d_{zz} orbital is obtained, whereas for a more balanced axial/rhombic ratio a tensor shape resembling a d_{xy} orbital is obtained.

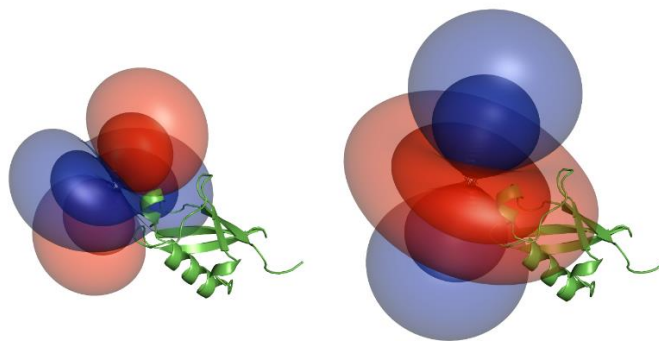


Fig. 3. Isosurfaces of the anisotropy parameters generated by the dysprosium (left) and thulium (right) complex of Ln-DOTA-M8-(4R4S)-SSPy when attached to ubiquitin S57C (inner layer: 1.0 ppm, outer layer: 0.25 ppm). While Dy³⁺ induces a tensor with a more balanced axial/rhombic ratio ($\Delta\chi_{ax}/\Delta\chi_{rh} = 2.1$), Tm³⁺ causes a tensor with a pronounced, high axial/rhombic ratio ($\Delta\chi_{ax}/\Delta\chi_{rh} = 6.5$). [23]

In order to study proteins by PCS NMR spectroscopy, HSQC experiments with a diamagnetic and a paramagnetic sample have to be performed. Based on the PCSs observed for the paramagnetic sample when compared to the diamagnetic sample, $\Delta\chi$ -parameters can be obtained by application of a Monte-Carlo simulation and the crystal structure can be refined in solution [103]. In most cases, the pseudocontact shift contribution for both dimensions of the spectrum is equal when expressed in units of ppm, therefore the shifted signals are found in a 45° angle relative to the corresponding diamagnetic signal. Due to the distance between the amide nitrogen and hydrogen atom of a common residue of ~ 1.0 Å [104], exceptions of this behaviour arise for N-H groups located on nodal planes of the anisotropy tensor or in regions with a very steep gradient. The obtained spectra are iteratively assigned and the obtained shifts fitted to an already solved X-ray structure or a Rosetta model.

PCSs give rise to multiple applications from structure refinement to localization of ligands [19, 20, 66-71]. Deviations from the back-calculated shift from the crystal structure can be used to determine flexible regions in the protein. Furthermore, a protein with a paramagnetic centre incorporated allows in combination with the Rosetta software package to perform more elaborate predictions of the protein structure [105]. The required input for the structure determination of a protein is then reduced to PCS data for all amino protons and the sequence. Based on the combination of a database containing structures for short oligopeptide fragments and the restraints given by the dataset of PCSs, a *de novo* structure determination of a biomacromolecule can be performed. Besides their use as structural restraints for the investigation of the structure of proteins, PCSs can also be exploited for the localization of ligands on a protein of interest [20, 66-71]. By determination of the $\Delta\chi$ -parameters from multiple tagging sites, the ligand bound to the protein can then be localized based on its PCS. This localization method was among other examples demonstrated for ¹H PCS of weakly bound ligands [69] and for ¹⁹F PCS in the intermediate exchange regime [66] as well as for strongly binding ligands [71].

2.3 Residual dipolar couplings

RDCs can be observed in NMR experiments, when partial alignment leads to an incomplete averaging of anisotropic magnetic interactions and a coupling with a dipole of the nuclear spins under investigation, e.g. the N-H bond dipole in proteins, leads to a splitting of signals [46]. Under isotropic conditions, the protein tumbling in solution results in an averaging to zero of the dipolar interactions. If dipolar couplings shall be observed, partial alignment is therefore required. This can be achieved by either introducing a paramagnetic centre covalently attached to the protein and thus leading to a field-dependent alignment or by the use of alignment media such as polyacrylamide or collagen gels leading to a field-independent alignment. Using RDCs, as in the case of PCSs, based on a known structure or a structure prediction, e.g. by Rosetta, the $\Delta\chi$ -parameters can be fitted and a refined structure will be obtained. Since in principle the orientation of all N-H bond vectors of the protein can be obtained by RDCs, the gained information is complementary to NOESY experiments and lead in combination with NOESY restraints to a more precise picture of the solution structure of the investigated biomacromolecule [106, 107]. In contrast to PCSs (R^{-3}) or PRE (R^{-6}), RDCs are independent of the lanthanide – spin pair distance.

$$\text{Dipolar coupling for two spins } i \text{ and } j \quad D_{ij} = k_{ij} \left[A_a (3 \cos^2 \theta - 1) + \frac{3}{2} A_r \sin^2 \theta \cos(2\phi) \right] \quad \text{Eq. (5)}$$

$$\text{with} \quad k_{ij} = -\frac{\gamma_i \gamma_j \mu_0 \hbar}{16\pi^3 r_{ij}^3} \quad \text{Eq. (6)}$$

DOTA-based LCTs, e.g. CLaNP-5 [24] or DOTA-M8-(8S)-SSPy [9], can yield such valuable structural restraints. As recently demonstrated by Mütener et al., using a DOTA-based tag bearing a reduction stable pyridine linker, RDCs up to 27.7 Hz were observed in in-vitro and in in-cell experiments [3]. These valuable orientation restraints were then used to perform a *de novo* structure determination for GB1 in combination with the Rosetta database. The obtained structure corresponds to the X-ray structure of the protein. For proteins without possibility to record an X-ray structure or when deviations for the structure in solution from the structure in crystal state are expected, RDCs present a valuable source of structural restraints.

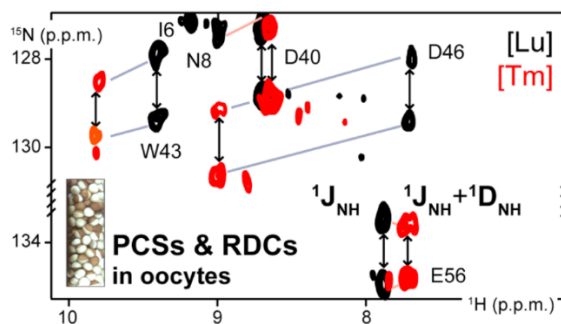


Fig. 4. RDCs measured on GB1 conjugated to an LCT in *Xenopus laevis* oocytes acquired at 600 MHz proton frequency [3]. (Reprinted with permission from T. Mütener, D. Häussinger, P. Selenko, F.-X. Theillet, *In-Cell Protein Structures from 2D NMR Experiments*, *J. Phys. Chem. Lett.*, 7 (2016) 2821-2825. Copyright 2016 American Chemical Society.)

RDCs are most valuable as a source of long-range structural restraints and can be used to give insights into specific biological problems, as e.g. the determination of the relative orientation or exact angle between two helices in proteins, the extent of bending for a long helix of a protein and the orientation of two subunits of a protein or protein-protein complex [108]. The combination of RDCs with other methods as X-ray crystallography can lead to highly refined solution structures of both proteins and nucleic acids, e.g. of a protein-protein complex for which the structures of the individual subunits have been determined previously by X-ray crystallography. A more exotic application of RDCs includes the determination of the stereochemistry of natural products [109]. Besides the use of RDCs for the structure determination of proteins, also the helix curvature, subunit or domain orientation and stoichiometry of homomultimeric nucleic acid complexes was investigated [110].

To conclude, RDCs provide valuable long-range structural restraints and are able to drive the field of structural biology to a new level of precision in the structure elucidation of biomacromolecules in solution.

3. Lanthanide chelating tags

3.1 Introduction & Overview

Paramagnetic NMR using lanthanides started with the exploitation of metal binding sites, particularly Ca^{2+} binding sites, since Ca^{2+} ions display a similar oxophilicity and ionic radius (112 pm [111]) when compared to Ln^{3+} ions (Ce-Lu: 114-98 pm [111]). Bertini et al. compared in a comprehensive study the full series of lanthanide ions in a similar coordination environment in the protein calbindin D_{9k} [112] (tensors of the full lanthanide series depicted in ref. [20]). Upon successful substitution of the calcium with a lanthanide ion, paramagnetic effects as e.g. PCS can be measured directly, as shown by Bertini et al. for Calbindin D_{9k} that possesses two calcium binding sites [113]. However, obviously this approach only works for proteins containing a metal binding site. In order to study other targets by paramagnetic NMR, artificial binding sites have to be introduced or attached to the protein of interest.

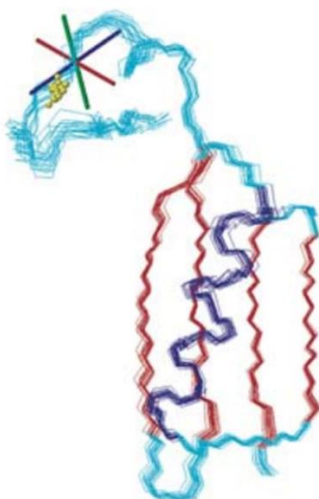


Fig. 5. Refined structure obtained using a two-point anchored LBP (metal position is indicated with yellow dots. The cross indicates the tensor orientation) [59].

Already in 2000, Ma et al. reported an EF-hand motif engineered to the N-terminus of a membrane-associated protein and showed thereby that the introduction of lanthanides into biomacromolecules without native metal binding site via LBPs and exploitation of the induced RDCs is possible [114]. LBPs mimic EF hand and zinc finger motifs in order to incorporate lanthanides into proteins with no natural metal binding site [60, 61, 114-117]. LBPs are advantageous in regard to tagging efficiency since peptide fusion always gives rise to very high ligation yields. However, a significant drawback is that the incorporation is only possible on the N- or the C-terminus of the investigated protein without severely altering the structure of the protein. Furthermore, labelling of the LBP leads to increased complexity of Protein NMR spectra. In order to gain more flexibility considering location of the attachment site, LBPs that are conjugated to cysteine residues were developed [118, 119], but often limited affinity to the lanthanides was problematic.

In order to circumvent the drawbacks associated with LBPs, synthetic LCTs were developed. Based on chelators as diethylenetriaminepentaacetic acid (DTPA) [120, 121] and ethylenediaminetetraacetic acid (EDTA) [21, 122], synthetically produced metal binding sites with high lanthanide affinity can be provided. The tags have to be attached to a single ligation site on the protein, whereas the thiol group of cysteine stands out as an ideal target due to the low abundance in protein sequences and the possibility for a chemoselective modification with LCTs. In order to provide a rigid tagging site in a well-defined region of the protein, typically cysteine residues in α -helices are chosen as attachment point. For the tagging of cysteine residues in proteins, suitable linkers had to be developed. The most practical systems are pyridine-2-thiol as well as methyl- and phenylsulphonyl activated linkers [1]. Unfortunately, the in biochemistry frequently used maleimides are not suitable for the attachment of LCTs to proteins because they yield an additional stereo-centre at the linkage site upon tagging reaction resulting in two diastereomeric species of the protein-tag construct. Recently, also artificial amino acids, e.g. genetically encoded *p*-azido-*L*-phenylalanine residues, which can be tagged in a bio-orthogonal fashion, were introduced by Otting et al [6]. Bio-orthogonal LCTs allow to investigate proteins that cannot be expressed in a stable form with a single cysteine on their surface. Furthermore, bio-orthogonal LCTs can be used to characterize proteins that bear multiple cysteine residues on their surface that are essential to their biological function. Despite the considerable effort to introduce bio-orthogonal LCT, this is definitively the most general approach and in many difficult cases of multi-cysteine proteins will be the only viable one.

3.2 Non-DOTA-based lanthanide chelating tags

Already in 2002, Dvoretzky et al. demonstrated the use of a pyridine-2-thiol activated linker moiety in combination with an EDTA-derived metal chelator for the application in paramagnetic NMR [123]. Upon ligation of the novel chelator, loaded among other paramagnetic metal ions with Yb^{3+} , to barnase H102C by formation of a disulphide linkage, the valuable PCS and RDC long-range structural restraints were analysed and fitted to the crystal structure of wild-type barnase.

The development of two-armed caged lanthanide NMR probe (CLaNP) LCTs was started in 2004, when Prudêncio et al. reported the first CLaNP, consisting of a DTPA moiety that is equipped with two reactive groups suitable for ligation to cysteine residues of the investigated protein [120]. Impressively, CLaNP gave rise to observable PCSs detected on nuclei farther away than 40 Å from the metal centre. Unfortunately, up to five shifted resonances were detected for each residue. However, the strong PCSs induced by CLaNP clearly already illustrated at this point in time the potential of the use of PCSs and RDCs in structural biology.

In 2004, Ikegami et al. developed a novel, enantiomerically pure EDTA-derived LCT that can be attached via a cysteine residue to a protein of interest [21]. The LCT was then used for a partial alignment of the protein in solution and RDCs up to a magnitude of 8 Hz were observed at a field strength of 18.8 T. Notably, the LCT was conjugated to the trigger factor, a ribosome-associated molecular chaperone, and an analysis of the tensors of the anisotropic magnetic susceptibility was performed. The study identified the drawback of EDTA-derived LCTs, i.e. the peak doubling due to

the generation of an additional stereo-centre on the nitrogen atom close to the introduced linker, which becomes chiral upon metal complexation.

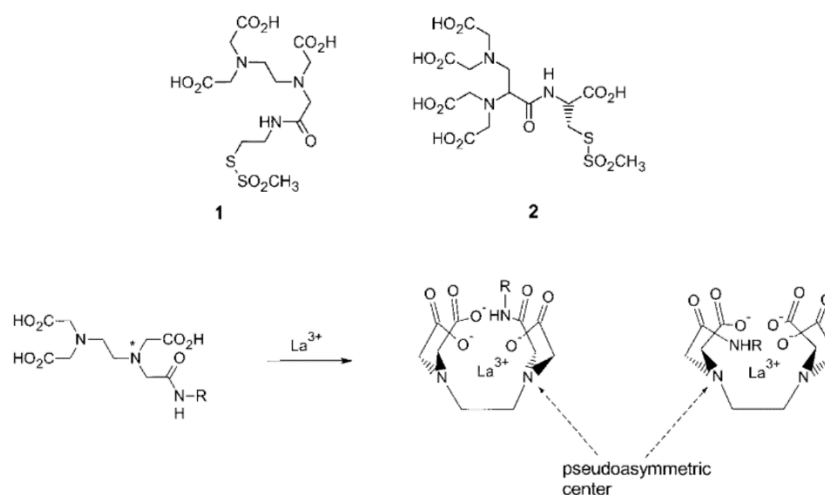


Fig. 6. Structure of EDTA-derived LCTs reported by Ikegami et al. [21] and Leonov et al. [62] While compound **1** yields an additional stereo-centre upon metal complexation due to formation of a pseudoasymmetric centre, compound **2** is designed with two identical substituents on both nitrogen atoms coordinating to the lanthanide ion in order to avoid this behaviour.

Leonov et al. developed in 2005 an improved version of the EDTA-derived LCT presented by Ikegami et al. [21], giving rise to only one set of signals when attached to the protein of interest [62]. The diastereochemically pure compound can be synthesized in 5 steps and the corresponding (*R,R*)- and (*S,R*)-diastereomers were separated by high-performance liquid chromatography (HPLC). Since each of the two nitrogen atoms, which are able to coordinate to the lanthanide ion, are substituted with two identical groups, neither of them becomes stereogenic upon metal complexation and only one set of signals is detected in the application of the LCT on the protein, thereby significantly facilitating analysis of the acquired spectra.

Haberz et al. extended the previous work by the synthesis of two new EDTA chelators in 2006 that were successfully attached to the trigger factor protein and apo-calmodulin [124]. Thereby the authors of the presented study could improve the precision of the solution structure of trigger factor and showcase the extremely high affinity of the new EDTA-derived LCT to lanthanides, exceeding the one observed for the natural metal binding site of apo-calmodulin. In 2008, Su et al. presented an LCT based on 4-mercaptomethyl-dipicolinic acid (4MMDPA) that can be readily attached to proteins via a disulphide linkage [125]. 4MMDPA loaded with Yb^{3+} yields PCSs up to 2 ppm when attached to a cysteine residue in position 68 of arginine repressor from *E. coli*. Notably, the LCT is immobilized on the proteins surface by an additionally coordinating carboxy group of the protein.

Man et al. developed in 2010 a novel small and rigid LCT, i.e. 3-mercapto-2,6-pyridinedicarboxylic acid (3MDPA) [126]. The applicability of the newly developed LCT was showcased by attachment to the N-terminal domain of the arginine repressor (ArgN) from *E. coli* and T4 lysozyme C54T C97A Q69C. The LCT yielded anisotropy parameters of $\Delta\chi_{\text{ax}} = 6.6 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 2.4 \times 10^{-32} \text{ m}^3$ for thulium loaded 3MDPA in the ArgN-3MDPA construct. While the simple

synthesis and the sufficient induced anisotropy parameters render the tag interesting for studies of protein structure and dynamics, possible interaction with the Glu21 sidechain of ArgN and the fact that only three donor atoms are used to coordinate the lanthanide ion constitute the drawback of the presented LCT when compared to rigid and high affinity ligands.

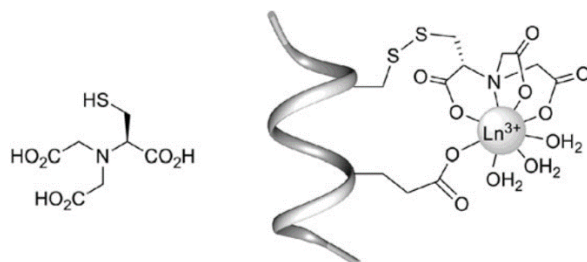


Fig. 7. Proposed lanthanide chelating mode of the NTA-derived LCTs in an α -helix with a carboxy sidechain in $i+4$ or $i-4$ position developed by Swarbrick et al. [127].

In 2011, Swarbrick et al. developed a small LCT that is attached to the α -helix of a protein of interest via a disulphide linkage and becomes rigidified directly on the protein by interaction of the lanthanide ion with an Asp residue in the $i+4$ position [128]. While the synthesis of the presented iminodiacetic acid-tag (IDA-SH) is extremely simple and attractive, i.e. three steps in an overall yield of 45%, the requirement of an Asp residue in the $i+4$ position restricts the applicability of the LCT exclusively to proteins, where a native or by mutation introduced cysteine is located in an α -helix with an Asp residue in the $i+4$ position. Swarbrick et al. further demonstrated that it is possible to use nitrilotriacetic acid-derived tags (NTA-SH) for the generation and analysis of large PCSs and RDCs [127]. Yagi et al. then investigated the optimal positioning of the IDA-SH and NTA-SH LCTs [129]. In a systematic study, in which the authors introduced the mentioned LCTs in an α -helix with either glutamate or aspartate residues in the positions $i-4$ or $i+4$, it was found that the largest anisotropy parameters can be observed for a NTA-SH LCT with a glutamate residue positioned at $i-4$. Furthermore, it was shown that the NTA-derived LCTs are able to produce sufficiently large PCSs independent of nearby carboxylate functional groups in the protein. In contrast to the NTA-SH LCT, the IDA-SH LCT requires an aspartate residue in the $i+4$ position.

Jia et al. presented in 2011 the use of the dipicolinic acid-derived tags, i.e. 4MDPA, for the generation of structural restraints induced by the paramagnetism of lanthanides [130]. The switch from 3MDPA to 4MDPA is highly favourable in order to suppress rotational averaging of the paramagnetic effects, i.e. the short and rigid tether as well as the symmetric assembly of the chelator and the lanthanide ensure a successful transfer of the paramagnetic effects from the lanthanide to the protein for 4MDPA. Further advantages are the small size of the tag that ensures no disturbances of the protein structure and the quantitative tagging yields. However, the anisotropy tensor parameters of $\Delta\chi_{ax} = -9.2 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = -5.4 \times 10^{-32} \text{ m}^3$ for Tm^{3+} on the ArgN construct and the possible interactions of DPA-derived LCTs with other residues of the protein [126] still offered room for further developments of LCTs.

In 2011, Peters et al. reported the synthesis of Cys-Ph-TAHA, a highly symmetric LCT based on the tris(2-aminoethyl)aminehexaacetic acid (TAHA) chelator [14]. The presented LCT induces PCSs up to 2 ppm and RDCs up to 18 Hz, while providing a metal affinity in the femtomolar range and only one set of signals in ^1H - ^{15}N HSQC spectra. Furthermore, the LCT shows a good thermal stability in the temperature range of 278-315 K. However, due to the carboxylic acid ligands, the LCT is pH-sensitive, i.e. the LCT cannot be used for example to monitor protein unfolding below pH 4.0. In order to show that the LCT is suitable for analysis of large proteins and their complexes, Cys-Ph-TAHA was successfully attached to a 90 kDa assembly of a Lac repressor protein in complex with DNA and isopropyl β -D-1-thiogalactopyranoside (IPTG).

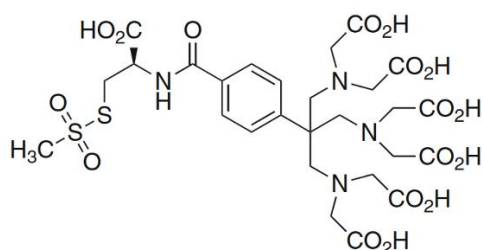


Fig. 8. Structure of the Cys-Ph-TAHA LCT [14].

An LCT ligated via a thiol-ene reaction to the protein was proposed in 2011 by Li et al [131]. The researchers attached the LCT incorporating a vinylpyridine moiety to ArgN and human ubiquitin and showed that the protein-tag conjugate is stable in presence of reducing agents as dithiothreitol (DTT) or 3,3',3''-phosphanetriyltripropanoic acid (TCEP) due to the inert thioether linkage. Tagging was achieved in quantitative manner using an eight-fold excess of the LCT and low molecular weight compounds could subsequently be removed by ion-exchange chromatography leading to the desired tag-protein conjugate in 80% yield. Sizeable anisotropy parameters of $\Delta\chi_{\text{ax}} = 13.8 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 7.5 \times 10^{-32} \text{ m}^3$ were induced on ubiquitin T22C using the Tm^{3+} -loaded LCT.

Yang et al. described in 2012 the synthesis of 4-vinyl(pyridine-2,6-diyl)bismethylenenitrilo tetrakis(acetic acid) (4VPyMTA) and its use as chelator in paramagnetic NMR [132]. The presented tag shows only one set of signals, the formed linkage of the LCT to the protein is stable under reducing conditions and the affinity for lanthanides is higher than the one of EDTA. The induced anisotropy is only in a medium range, i.e. $\Delta\chi_{\text{ax}} = -2.3 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = -0.5 \times 10^{-32} \text{ m}^3$ for Tm^{3+} on ubiquitin E64C, however this leads to two advantages: i) convenient assignment of ^1H - ^{15}N HSQC spectra, ii) a large number of PCS that can be detected. The development of reduction-stable linker moieties is highly desirable, however, the ligation yield of 80% still left room for improvement in further developments.

As a further development among the earlier introduced dipicolinic acid LCTs, Huang et al. reported the synthesis and application of a related 4'-mercapto-2,2': 6',2''-terpyridine-6,6''-dicarboxylic acid (4MTDA) LCT [133]. Upon attachment of the LCT to the protein of interest, interactions of a carboxylate sidechain of the protein with free coordination sites of the lanthanide in the chelator immobilize the LCT on the proteins surface. When attached to the protein, the presented LCT yields large PCSs and RDCs up to 13 Hz. Furthermore, the described LCT can also be

used for fluorescence and EPR studies. Unfortunately, 4MTDA exhibits inferior reactivity than the related, previously published 4MMDPA [125] and 3MDPA [126] LCTs, resulting in ligation yields of only 50-70%.

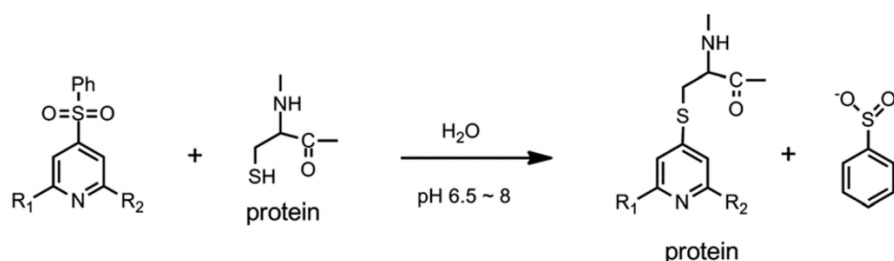


Fig. 9. Tagging reaction of the 4PS-DPA and 4PS-PyMTA LCTs bearing a phenylsulphonyl moiety reported by Yang et al [134].

Yang et al. described two LCTs with a phenylsulphonyl leaving group (4-(phenylsulphonyl)-pyridine-2,6-dicarboxylic acid (4PS-DPA) and 4-phenylsulphonyl-(pyridin-2,6-diyl)bismethylenenitrilo tetrakis(acetic acid) (4PS-PyMTA)) leading to an extremely short and rigid thioether linkage with the protein of interest [134]. Tm^{3+} loaded 4PS-PyMTA leads to large observed anisotropy parameters of $\Delta\chi_{\text{ax}} = 16.6 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 3.9 \times 10^{-32} \text{ m}^3$ when attached to ubiquitin G47C R72A R74A. Due to the generation of reduction-stable protein-LCT linkages, the authors anticipated that the phenylsulphonylpyridine moiety finds broad application in structural biology and related fields.

In 2015, Loh et al. presented clickable IDA and NTA tags bearing an alkyne linker moiety that can be ligated in a copper catalysed azide alkyne click reaction to a site-specifically introduced *p*-azido-*L*-phenylalanine (AzF) residue in ubiquitin and GB1 followed by an analysis of paramagnetic effects and anisotropy parameters [135]. For the thulium loaded alkyne-NTA LCT, maximal anisotropy parameters of $\Delta\chi_{\text{ax}} = 16.3 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 7.8 \times 10^{-32} \text{ m}^3$ were observed upon ligation to the GB1 V21 construct. The authors suggest to use 150 mM NaCl during the ligation reaction, which proved to be beneficial for the AzF mutants of GB1, ERp29, PpiB, the intracellular domain of the p75 neurotrophin receptor and the West Nile virus NS2B-NS3 protease in terms of precipitation of the protein during the attachment of the LCT.

In line with the findings of Loh et al., Jiang et al. developed diethylene-triamine-tetraacetate propyl-1-yne (DTTA-C3-yne) and diethylene-triamine-tetraacetate butyl-1-yne (DTTA-C4-yne) LCTs that were attached in bio-orthogonal fashion to ubiquitin and the enzyme E II B using a copper catalysed azide-alkyne click reaction [136]. Upon attachment to ubiquitin incorporating a *p*-azido-*L*-phenylalanine (AzF) residue in position 18, the Tb^{3+} and Tm^{3+} loaded LCTs yielded sizeable anisotropy parameters for the analysed PCSs and delivered RDCs up to 8.3 Hz. Furthermore, the authors presented a refined model of the interactions in the UBA1-ubiquitin complex based on intermolecular PCSs. Chen et al. reported in 2017 the use of DTPA-derived LCTs, i.e. 4PS-PyDTTA and its methyl-substituted analogue 4PS-6M-PyDTTA [137]. Both of the investigated LCTs yield large PCSs and RDCs, more specifically, the methyl-substituted LCT generally shows larger anisotropy parameters on ubiquitin and SrtA with the thulium loaded LCT, most likely due to suppressed rotational averaging caused by the additional methyl group. The presented LCT offers a high stability

of the formed complexes - neither 1.0 eq. of EDTA in the ubiquitin sample, nor the naturally present calcium binding motif in SrtA was able to interfere with the binding of the lanthanide ions by the DTPA chelator.

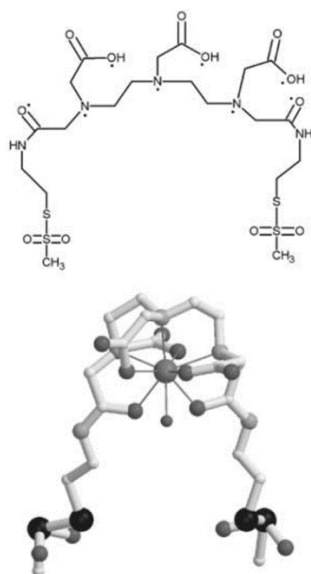


Fig. 10. Structure of the DTPA-based two-point anchored LCT CLaNP (dots: coordinating atoms) [120].

The developed EDTA-derived LCTs can lead to peak doubling observed for the protein in HSQC experiments and therefore to two sets of signals with two associated χ -tensors, since some of the EDTA ligands become chiral upon metal complexation and lead to two diastereomeric species in the protein-tag conjugate. Furthermore, the obtained PCSs were only of modest size. Using DTPA-based tags, up to eight different sets of signals can be expected and for two-armed lanthanide probes, namely CLaNP, up to five different sets of signals were observed in practice [120]. In order to produce LCTs with a higher affinity of the ligand for the lanthanide ion, increased PCSs observed in HSQC experiments, more rigid chelation of the lanthanide ion, one set of signals, ability to preload the LCT with metal ions avoiding the titration of the protein with lanthanide ions until complete loading is achieved, the introduction of DOTA-based LCTs for paramagnetic NMR spectroscopy was accomplished.

3.3 DOTA-based lanthanide chelating tags

DOTA provides an ideal framework for the chelation of lanthanide ions and development of LCTs. Besides the fact that DOTA complexes of lanthanide ions display a very high thermodynamic stability (equilibrium constant for the formation of Gd(III)-DOTA = 10^{25} M^{-1} at pH 7 [138]), they can also form neutral lanthanide complexes, when suitably modified. Due to the excellent properties of DOTA as ligand in Gd³⁺ complexes, the chelator has found widespread applications in MRI as well as in radiopharmacy [139, 140].

The first DOTA-based LCT was reported by Häussinger et al. in a conference abstract in the year 2005 [141]. Based on the DOTA framework that has a binding constant that is 10 orders of magnitude larger when compared to EDTA,

a lanthanide tag equipped with a short $\text{SCH}_2\text{CH}_2\text{NH}$ linker between the cysteine sulphur of the protein and the carboxy group of the DOTA framework was developed and applied to characterize the 50 kDa dimer of the calcium dependent cell adhesion protein ECAD12. This DOTA-based LCT can be used over a wide pH range and in the presence of EDTA and phosphate buffer.

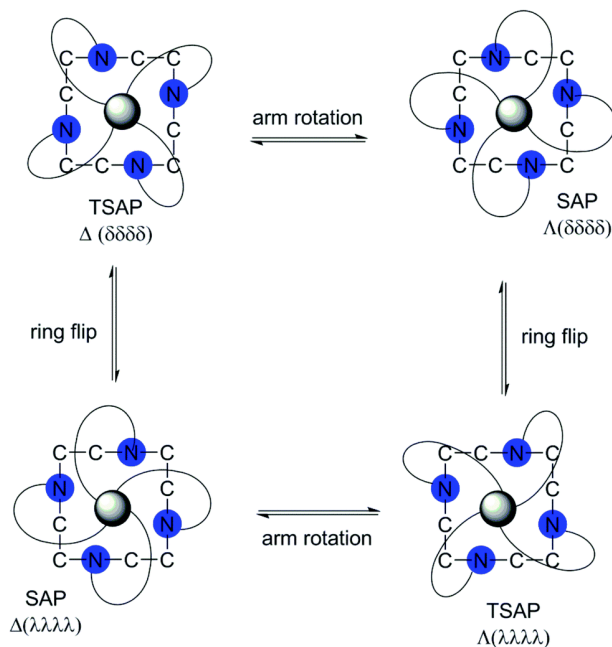


Fig. 11. Conformations of DOTA-derived ligands upon metal coordination [142].

However, the complex conformational mobility of the DOTA ligand, i.e. the two different conformations of the basic ring scaffold ($\delta\delta\delta\delta$ and $\lambda\lambda\lambda\lambda$) and the two different conformations of the pendant acetate arms of the chelator (Δ and Λ), lead, upon attachment to the protein of interest, in theory to four different conformational isomers ($\Delta(\delta\delta\delta\delta)$, $\Delta(\lambda\lambda\lambda\lambda)$, $\Lambda(\delta\delta\delta\delta)$, $\Lambda(\lambda\lambda\lambda\lambda)$). For certain residues in ubiquitin S57C labelled with the Dy-DOTA-SSPy LCT up to three different resonances were observed in ^1H - ^{15}N HSQC spectra. Furthermore, it is reasonable to suspect that the ring flexibility of Ln-DOTA-SSPy can lead to a significant motional averaging of the PCS in the case of fast conformational exchange.

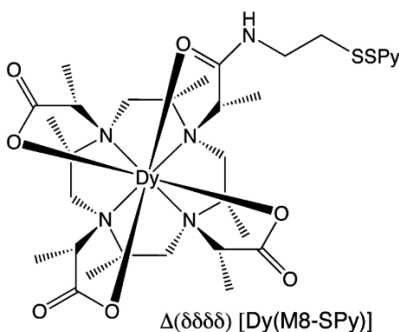


Fig. 12. Structure of the rigidified DOTA-based Dy-DOTA-M8-(8S)-SSPy in TSAP conformation [9]. (Reprinted with permission from D. Häussinger, J.-R. Huang, S. Grzesiek, DOTA-M8: An Extremely Rigid, High-Affinity Lanthanide Chelating Tag for PCS NMR Spectroscopy, *JACS*, 131 (2009) 14761-14767. Copyright 2009 American Chemical Society.)

In order to decrease the ring motions of the DOTA-derived LCT resulting in motional averaging of PCS, methyl substituents were introduced both on the side arms and the macrocyclic polyamine scaffold [9]. The methyl group substituents on the nitrogen containing macrocycle adopt an equatorial-upper position [121, 142] when the ligand is coordinated to a lanthanide metal ion and prevent motional averaging and line broadening of signals by locking the 12-membered ring in a $(\delta\delta\delta\delta)$ conformation [9].

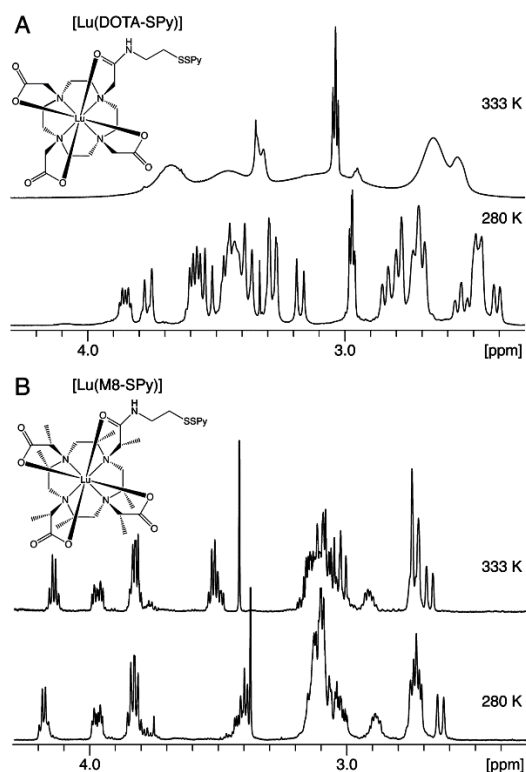


Fig. 13. A) Motional averaging of NMR signals using a non-rigidifying DOTA chelator. At 280 K, the motions are clearly decreased. B) For the methyl-substituted chelators, even at elevated temperatures the motions are severely decreased when compared to the non-methyl-substituted analogue. All spectra were acquired at 600 MHz proton frequency [9]. (Reprinted with permission from D. Häussinger, J.-R. Huang, S. Grzesiek, DOTA-M8: An Extremely Rigid, High-Affinity Lanthanide Chelating Tag for PCS NMR Spectroscopy, *JACS*, 131 (2009) 14761-14767. Copyright 2009 American Chemical Society.)

The strongly paramagnetic, extremely rigidified LCT displayed very large PCSs of over 5 ppm and RDCs over 20 Hz when attached to ubiquitin S57C. Besides the fact that these properties were unprecedented for a single-point attached LCT, the tag can be used under extreme buffer conditions, due to the high stability of the lanthanide complex over a wide pH range and very high affinity of the chelator towards lanthanide ions.

Interestingly, Dy-DOTA-M8-(8S)-SSPy exhibits two sets of signals when attached to a protein of interest resulting in a complicated analysis of ^1H - ^{15}N HSQC spectra containing the PCSs and RDCs and hence the structural restraints for the investigation of the structure of the protein. Opina et al. reported that the two sets of signals originate in the presence of two different conformational species of the LCT due to repulsive interactions of two methyl groups, so that the two conformers are interconverted to each other by rotation of the side arms of the Ln-DOTA-M8-(8S)-SSPy LCT on a timescale of approximately 4 h [142]. Opina et al. further showed that the two conformational species of the Ln-DOTA-M8-(8S)-SSPy LCTs can be observed in significant amounts for members of the lanthanide series between Eu^{3+} and Yb^{3+} and that the observed ratio of the two conformational species is governed by the ionic radius of the chelated lanthanide ion [142]. Strickland et al. investigated the temperature and pH-sensitivity of paramagnetic susceptibility tensor of Ln-DOTA-M8-(8S)-SSPy LCTs [28]. The presented study investigating different lanthanide complexes showed that the temperature sensitivity depends on the size of the lanthanide ion and the conformational properties of the ligand. More specifically, a larger lanthanide ion as e.g. Dy^{3+} leads to coordination of an apical water molecule and higher sensitivity towards temperature changes, while for smaller lanthanides as e.g. Yb^{3+} , ^{17}O NMR experiments showed that no water molecule is coordinated to the LCT and hence the induced anisotropy does not significantly vary in a temperature range of 288-308 K [28]. The effect on the paramagnetic susceptibility tensor in a pH range of 5.0-7.4 was shown to be not significant [28]. Since the publication of the Dy-DOTA-M8-(8S)-SSPy LCT in 2009, it was applied in a range of projects. Among others, the investigated examples include the localization of inhibitors on proteins [71], the exploration of the metal position of mobile spin labels with new fitting routines [143] as well as studies of extremely large protein complexes [12].

Joss et al. reported that the Ln-DOTA-M8-(4R4S)-SSPy stereoisomer of the Ln-DOTA-M8-(8S)-SSPy LCT shows only one conformational species for the whole lanthanide series [23]. Furthermore, the paramagnetic susceptibility of the lanthanide chelated by the ligand is significantly altered when compared to its predecessor Ln-DOTA-M8-(8S)-SSPy, a result that provides evidence for the influence of the coordination polyhedron and the ligand-metal interactions on the magnetic anisotropy as it was proposed earlier by Mironov et al [144].

Besides the inversion of the stereochemistry of the eight-fold methyl substituted LCT, introducing bulky and large groups on the side arms in a LCT can also lead to suppression of the conformational interchanges of the pendant arms and basic ring scaffold. Graham et al. presented a DOTA-amide LCT inducing large PCSs and exhibiting only one conformational species and hence set of signals [63]. The increased anisotropy transferred from the LCT to the protein when compared to other non-methyl substituted LCTs originates most likely in the strongly suppressed rotation of the LCT on the proteins surface due to the large and sterically demanding side arms, a result that is

supported by the surprisingly small intrinsic shift range of 170.5 ppm in ^1H spectra of the ytterbium complex when compared to the anisotropy induced on the protein.

Lee et al. presented a DOTA-derived, compact and hydrophilic LCT with a significantly shorter disulphide linker when compared to the DOTA-M8-(8S)-SSPy LCT that can be synthesized in 8% overall yield [25]. Although the LCT only includes methyl substituents on the pendant arms and none on the basic ring scaffold, the presented LCT leads to observation of large anisotropy parameters ($\Delta\chi_{\text{ax}} = 26.6 \times 10^{-32} \text{ m}^3$, $\Delta\chi_{\text{rh}} = 6.0 \times 10^{-32} \text{ m}^3$, for $\text{Ln} = \text{Dy}^{3+}$) on ubiquitin A28C. The results obtained by Lee et al. show in an impressive manner the large influence of the tagging site on the observed anisotropy parameters. While for ubiquitin A28C anisotropy parameters of $\Delta\chi_{\text{ax}} = -19.4 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = -7.8 \times 10^{-32} \text{ m}^3$ for the thulium loaded LCT are observed, on the 20 kDa-sized protein construct HPPK S112C C80A anisotropy parameters of $\Delta\chi_{\text{ax}} = 54.5 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 12.5 \times 10^{-32} \text{ m}^3$ were reported. The presented LCT constitutes, due to the large induced anisotropies and the convenient synthesis, a highly promising LCT for a diverse range of applications.

Based on the presented chelators by Lee et al. [25], Lee et al. developed the corresponding series of two-armed LCTs [145]. As expected for the switch from a single-armed LCT to a two-armed LCT, the anisotropy parameters induced by the two-armed LCTs are significantly improved by a factor of ~ 2.5 (anisotropy parameters on ubiquitin E24C A28C: $\Delta\chi_{\text{ax}} = 47.1 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{\text{rh}} = 18.6 \times 10^{-32} \text{ m}^3$, for $\text{Ln} = \text{Tm}^{3+}$). PCSs induced by LCTs bearing hydroxyl donor groups were reported to depend on pH [25, 146]. However, Lee et al. observed for the mentioned two-armed LCT only a minor pH-sensitivity.

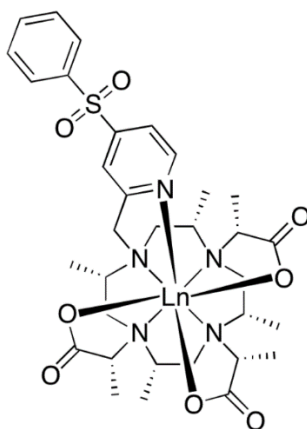


Fig. 14. Structure of Ln-DOTA-(3R4S)-M7Py suitable for in-cell NMR in oocytes reported by Müntener et al. [3]. (Reprinted with permission from T. Müntener, D. Häussinger, P. Selenko, F.-X. Theillet, *In-Cell Protein Structures from 2D NMR Experiments*, *J. Phys. Chem. Lett.*, 7 (2016) 2821-2825. Copyright 2016 American Chemical Society.)

In order to provide a reduction-stable linkage suitable for in-cell applications of the DOTA-M8 LCT, Müntener et al. used a pyridine derived linker that reacts with the cysteine residue of the protein to give a stable thioether linkage (Ln-DOTA-(3R4S)-M7Py) [3]. In order to improve the tagging reaction, an additional fluorine substituent was introduced on the pyridine in the ortho position of the leaving group, thereby stabilizing the Meisenheimer-Complex

formed upon reaction of the pyridine linker of the LCT with the cysteine residue of the protein [147]. The observed magnetic anisotropy parameters for Ln-DOTA-(3R4S)-M7FPy are similar to Ln-DOTA-M8-(4R4S) [23]. This result can be explained by an improvement of the obtained paramagnetic effects due to the shorter and more rigid linker, but a simultaneous decrease due to the unfavourable rotation of the LCT on the proteins surface caused by the less optimal linkage angle [147]. A similar approach was published by Pan et al. shortly after, illuminating the potential of this method for the investigation of protein structures in living cells [2].

In order to attach DOTA-derived LCTs to unnatural amino acids, Loh et al. presented two LCTs that can yield tag-protein conjugates by a copper catalysed azide-alkyne cycloaddition to a *p*-azido-*L*-phenylalanine residue [6]. The described LCTs induce reasonable PCS on ubiquitin Glu18AzF constructs and the bio-orthogonality of the linkage presents also an interesting approach for labelling of other biomacromolecules than proteins. However, the sample should tolerate copper ions, e.g. His₆ tags inhibit the copper-alkyne click reaction and hence formation of the tag-protein linkage.

A further approach towards the applicability of the Ln-DOTA-M8 LCTs for measurement of in-cell NMR constitutes the introduction of a carbamidemethyl linker bearing an iodide leaving group on the acetyl amide moiety that is reactive in a selective fashion towards thiol nucleophiles [4]. Hikone et al. modified the Ln-DOTA-M8-(8S)-SSPy LCT accordingly (Ln-M8-CAM-I) and demonstrated the PCS analysis of M8-CAM-tagged ubiquitin mutants that were incorporated in HeLa cells by electroporation [4]. The stable thioether linkage formed upon tagging enables the application of the tag under reductive conditions. However, when compared to the original Ln-DOTA-M8-(8S)-SSPy LCT, the anisotropy is severely decreased by the significantly longer, flexible linker part.

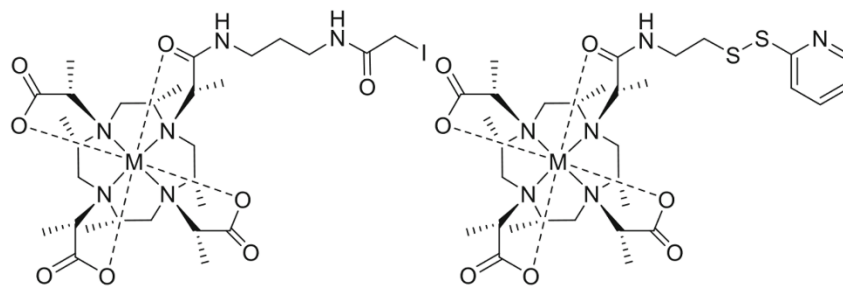


Fig. 15. Comparison of the structure of Ln-DOTA-M8-CAM-I reported by Hikone et al.[4] and Ln-DOTA-M8-SSPy reported by Häussinger et al. [9].

As a next step in the development of the DOTA-M8 tags, an LCT with isopropyl substituents on the basic ring scaffold (Ln-P4M4-DOTA) was synthesized and tested with several protein mutants [148]. Since the 1D spectrum shows a similar shift range as the one of DOTA-M8-(4R4S)-SSPy, the donor atoms were shown by DFT calculations to have a positional deviation of only 0.11 Å when compared to its eight-fold methyl-substituted predecessor and the metal centre mobility was shown by the proposed methodology of Suturina et al. [143, 149] to correspond to its predecessor, it can be concluded that the enhancement of the PCSs by a factor of almost 2 can be most likely

attributed to a more hindered rotation of the LCT on the proteins surface caused by the isopropyl substituents on the basic ring scaffold.

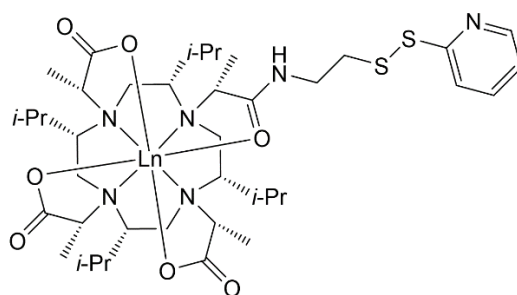


Fig. 16. Structure of a lanthanide complex of P4M4-DOTA in SAP conformation [148].

Besides the investigation of isopropyl substituents by Joss et al., a new reductively stable pyridinethiazole linker moiety was proposed by Müntener et al. and attached to the seven-fold methyl-substituted LCT to obtain the Ln-M7PyThiazole-SO₂Me-DOTA LCT [147]. It was shown that the enhanced rigidity of the linker moiety, when compared to the rather flexible LCT-CONH-CH₂-CH₂-S-Cys^{prot} linkage, and the more favourable attachment angle, which leads to decreased rotational averaging of the induced paramagnetic effects, as well as the improved synthetic access provide an extremely favourable LCT in terms of synthetic access, stability of the tag-protein linkage under reductive conditions as well as magnitude of observed PCSs and RDCs [147].

Clearly, it was obvious to test the newly developed linker moiety on an LCT equipped with an isopropyl-substituted backbone. The synthesized Ln-P4T-DOTA LCT shows large magnetic anisotropies induced by the different lanthanide complexes upon attachment to ubiquitin S57C, ubiquitin K48C and human carbonic anhydrase II (hCA II) S166C C206S and hence induces strong paramagnetic effects, i.e. PCS and RDC, in the tag-protein conjugates [150]. Furthermore, the large paramagnetic effects could also be used under reductive conditions, e.g. to obtain *de novo* derived in-cell protein structures from 2D NMR experiments [3].

The currently largest anisotropies induced by DOTA-derived LCTs are induced by a single-armed phenylsulphonated pyridine derivative of DOTA investigated by Yang et al. that constitutes a rigid and stable paramagnetic tag in analysis of protein structures and produces very large PCSs [151]. However, no conversion was observed in the tagging reaction with pH 7.8 or below and the lanthanide complex could only be attached to the protein using harsh conditions for protein tagging, i.e. pH 9.0 over 16 h or pH 9.3 over 18 h, hampering the general applicability of this newly developed LCT. While the extremely large PCSs can result in a cumbersome assignment process of the paramagnetic signals for small- to medium-sized proteins, the presented LCT would be very suitable for applications on large-sized proteins as well as the localization of ligands within proteins over large distances.

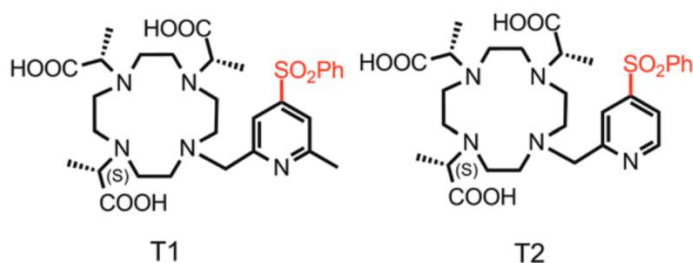


Fig. 17. Structures of the DOTA-derived LCTs reported by Yang et al. [151] The additional methyl substituent in T1 leads to a decrease of rotational and motional averaging and thus rise to the extremely large PCSs observed.

Besides introduction of more steric bulk on the basic ring scaffold and more rigid linkers causing less motional averaging of the paramagnetic effects, the LCT presented by Keizers et al. (CLaNP-5) that bears two linker arms constitutes an interesting approach [24, 152]. Although the LCTs equipped with two attachment sites to the protein require more challenging design and engineering of protein double mutants, which contain two cysteine residues in an optimal distance for the formation of the two disulphide linkages, the LCT presents an important advantage for its application in the investigation of ligand-protein interactions when compared to single-armed LCTs, i.e. the induced paramagnetic effects are predictable and independent of the environment of the LCT on the protein surface. The prediction can be used for a low-resolution model of the localization of a weakly binding ligand within a protein by PCS, removing the need for isotope labelling of both the ligand and the protein [69]. The initial low resolution model of the positioning of the ligand can give important structural insights on protein-fragment hit complexes in fragment-based drug discovery studies. When one wishes to localize strongly binding ligands, ^{19}F PCS and a full assignment of ^1H - ^{15}N HSQC spectra of paramagnetic protein-tag-ligand conjugates can provide the position of the ligand in a high accuracy [71].

An earlier developed DOTA-derived CLaNP LCT constitutes the CLaNP-3 LCT reported in 2007 by Vlasie et al. [153]. The two-armed LCT exhibits multiple isomers that are in exchange, but to the delight of the researchers only one signal was observed for the majority of the amides in ^1H - ^{15}N HSQC experiments, which paved the way for the determination of one single set of parameters of the induced anisotropy. By using the Yb^{3+} loaded LCT, PCSs were observed over a maximal distance of 35 Å away from the metal centre.

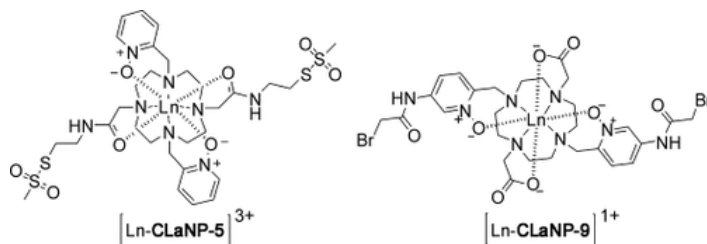


Fig. 18. Structural comparison of CLaNP-5 [24] and CLaNP-9 [154], which incorporates two carboxy groups that decrease the overall charge of the lanthanide complex.

In order to reduce the rather high, three-fold positive charge of the CLaNP-5 LCTs that can possibly lead to interactions with the protein, Liu et al. substituted the N-oxide moieties by nitro-substituted phenol groups, giving rise to an LCT containing two negatively charged alkoxides upon coordination of the three-fold positively charged metal centre thus resulting in the single positively charged CLaNP-7 LCT [155]. For the CLaNP-7 LCT, the orientation of the magnetic susceptibility tensor is pH-dependent when a histidine residue is located in the immediate vicinity of the binding site, a feature that could originate in interactions of the histidine residue with an apically coordinated water/hydroxyl molecule. In a further study, Liu et al. presented the CLaNP-9 LCT that delivers large PCSs on a Paz-Cu^{II} E51C E54C construct and possesses an overall charge of +1 as CLaNP-7 [154].

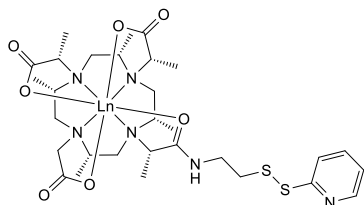
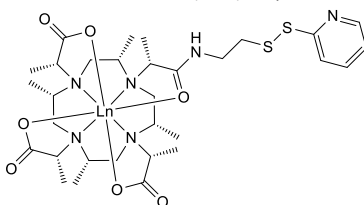
Lee et al. reported in 2016 an enantiomeric pair of pyridine-2-thiol activated, two-armed LCTs that gives rise to two different alignments of the magnetic susceptibility tensor [146]. For the thulium loaded LCT, large RDCs of up to 24.8 Hz and large PCSs were observed on an ubiquitin E24C A28C double mutant. Interestingly, the observed PCSs and associated anisotropy parameters at pH 6.5 are 40% smaller than those obtained at a pH of 8.0. Since the two LCTs can be readily synthesized in parallel and two different magnetic susceptibility tensors can be obtained using the same tagging site, the presented LCTs can be conveniently used for studies of protein structure and ligand binding.

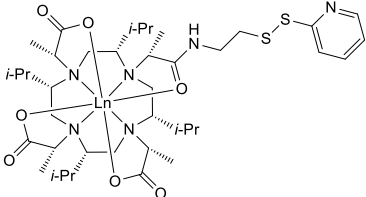
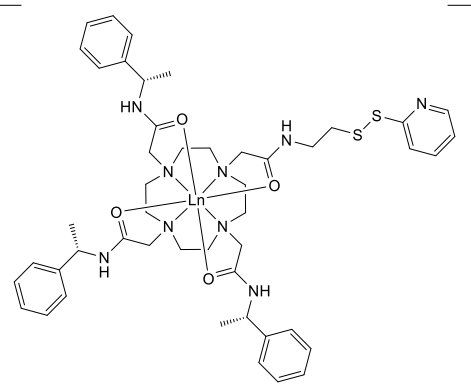
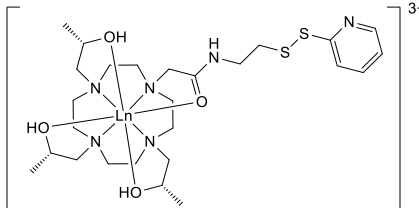
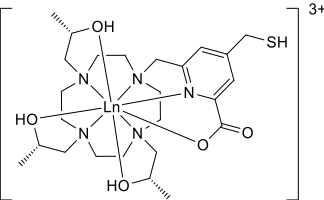
Besides disulphide, pyridine-derived and iodoacetamide linkers, bromoacetamides were shown by Wu et al. to be of significant value for ligation of an LCT to a phosphorothioate group in DNA [30]. The reduction-stable linkage that is formed between the LCT equipped with a bromo leaving group and the phosphorothioate offered the possibility for the authors to detect PCSs on DNA and subsequently fit the corresponding anisotropic $\Delta\chi$ -parameters. DOTA-derived LCTs are ideal scaffolds for the development of LCTs that can be attached to nucleic acids, since the strong binding of the lanthanide ion by the chelator and the inertia of the synthesized complexes ensures that no interferences between the metal ion and the phosphate backbone of the nucleic acid occur. The stereogenic phosphorus atom generated by the introduction of a phosphorothioate group leads to different sets of observed PCS. However, it is possible to separate the diastereomers by reverse-phase HPLC on a semipreparative C18 column leading to only one set of signals observed [30], thus opening the avenue of the transfer of the PCS methodology from proteins to DNA and RNA.

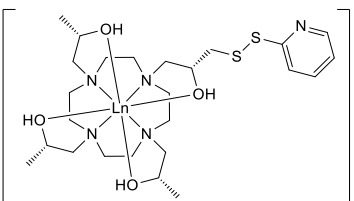
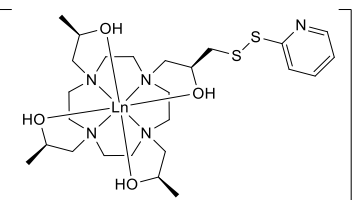
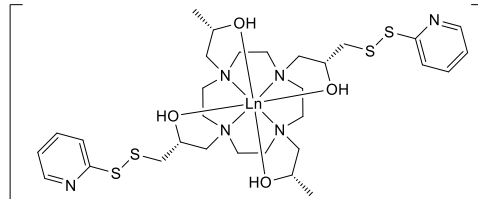
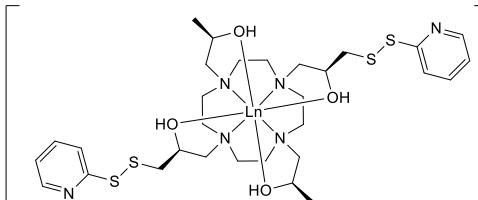
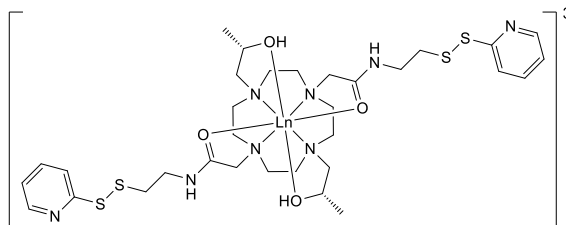
3.4 Tabular overview of DOTA-derived LCTs

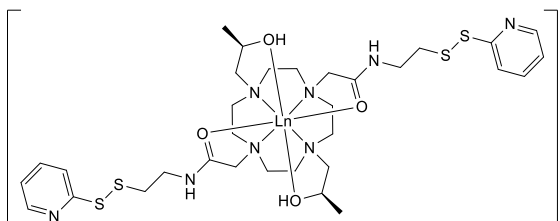
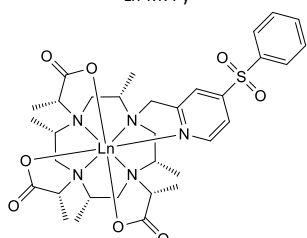
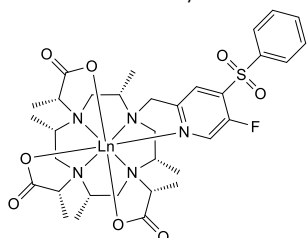
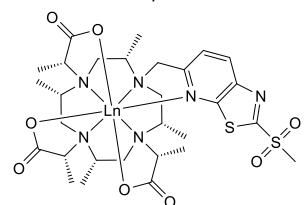
In the following, all DOTA-derived LCTs are listed with the reported anisotropy parameters and the corresponding tagging conditions. It is certainly clear, that the reported $\Delta\chi$ -values suffer from large variability, as they are usually obtained by the single-point approximation which strictly in no case holds true. In particular, pronounced mobility can lead to over-estimated $\Delta\chi$ -values in combination with elongated metal – attachment site distances [143, 149, 156]. But nevertheless the $\Delta\chi$ -values, especially when obtained for several different proteins, allow a rough estimation of the performance of a LCT – when taken with a pinch of salt. It is noteworthy, that the extent of anisotropy within the lanthanide metal series for a given LCT varies dramatically with minute differences of the coordination polyhedron geometry and donor atom set – in many cases Dy causes larger anisotropy than e.g. Tm, but there are also several examples with the reversed order.

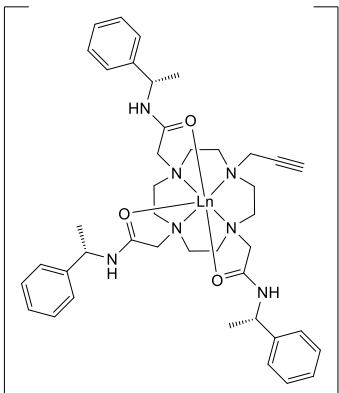
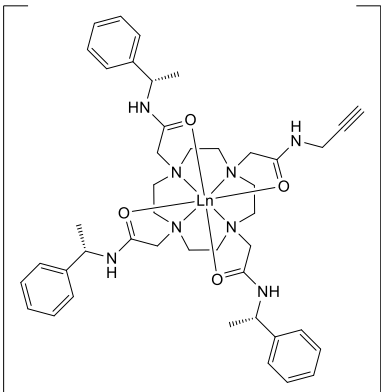
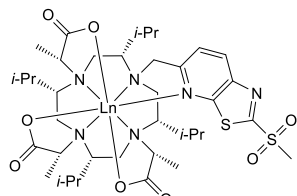
Table 1: Overview on DOTA-derived LCTs including tagged protein constructs, PCS- and RDC-derived observed anisotropy parameters, Ln-C_B distance as well as the used conditions for the tagging reaction.

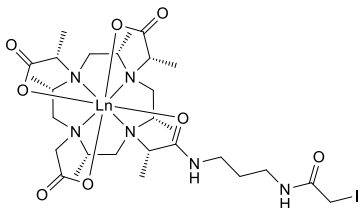
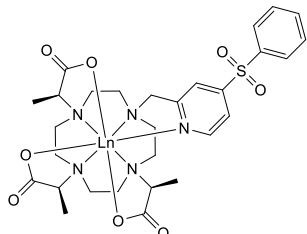
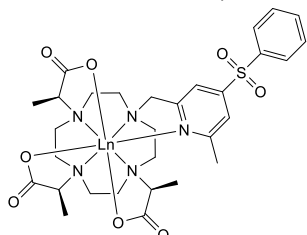
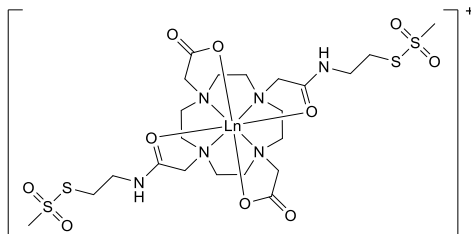
name & structure	protein construct	Ln ³⁺	$\Delta\chi_{ax, PCS}$ [10 ⁻³² m ³]	$\Delta\chi_{rh, PCS}$ [10 ⁻³² m ³]	$\Delta\chi_{ax, RDC}$ [10 ⁻³² m ³]	$\Delta\chi_{rh, RDC}$ [10 ⁻³² m ³]	distance Ln-C _B [Å]	tagging conditions & remarks	ref.
Ln-DOTA-M8-(8S)-SSPy 	Ubi S57C (PDB 1D3Z)	Dy ³⁺	28	17	21	6	7.9	3.0 eq. LCT, pH 5.0, 10 mM phosphate, rt, 3 h	[9]
	Ubi K6C (PDB 1D3Z)	Dy ³⁺	44	19	22	12	9.2		
Ln-DOTA-M8-(4R4S)-SSPy 	Ubi S57C (PDB 1UBI)	Dy ³⁺	-8.7	-4.1	-	-	7.2	3.0 eq. LCT, pH 7.0, 10 mM phosphate, rt, 3 h	[23]
		Tm ³⁺	19.6	3.0	-	-			
		Tb ³⁺	-8.2	-3.3	-	-			
	hCA II S50C C206S (PDB 3KS3)	Dy ³⁺	-21.0	-11.1	-	-	7.6	3.0 eq. LCT, pH 6.8, 10 mM phosphate, rt, 3 h	[148]
		Tm ³⁺	36.1	11.5	-	-			
		Tb ³⁺	-16.7	-11.8	-	-			
hCA II S166C C206S (PDB 3KS3)	Dy ³⁺	-21.2	-4.5	-	-	6.6	6.0 eq. LCT, pH 6.8, 10 mM phosphate, rt, overnight	[148]	
	Tm ³⁺	29.7	7.6	-	-	5.9			
Ln-P4M4-DOTA	Ubi S57C (PDB 1UBI)	Dy ³⁺	-16.7	-7.5	14.1	-4.8	8.7	6.0 eq. LCT, pH 7.0, 10 mM phosphate, rt, overnight	[148]
		Tm ³⁺	26.0	1.0	20.0	7.3	7.0		
	hCA II S166C C206S (PDB 3KS3)	Dy ³⁺	-29.7	-11.5	-	-	7.0	6.0 eq. LCT, pH 6.8, 10 mM phosphate, rt, overnight	
		Tm ³⁺	41.9	17.1	44.1	-11.1	7.2		

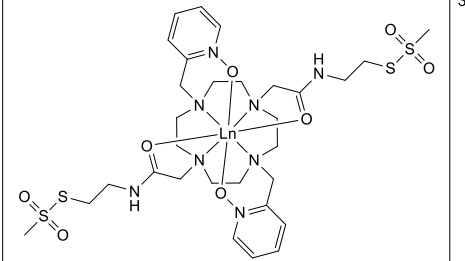
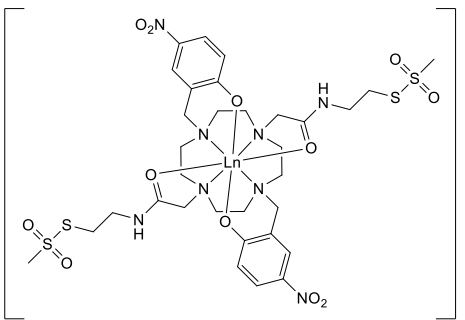
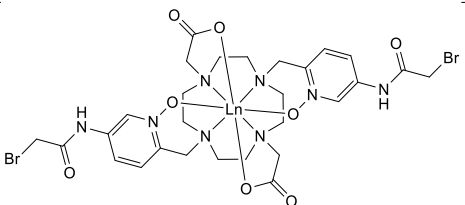
																
<p>Ln-C1</p> 	ArgN C68 (PDB 1AOY)	Dy ³⁺	-29	-11	-	-			8.7							
		Tb ³⁺	-27	-4	-	-										
		Tm ³⁺	37	12	-	-										
		Yb ³⁺	13	3	-	-										
	Ubi A28C (PDB 1D3Z)	Dy ³⁺	-10.8	-3.5	-	-			5.8							
		Tb ³⁺	-9.0	-2.3	-	-										
		Tm ³⁺	7.1	1.6	-	-										
		Yb ³⁺	2.4	0.5	-	-										
<p>Ln-C5</p> 	Ubi A28C (PDB 2MJB)	Dy ³⁺	8.2	5.3	-	-			7.5							
		Tb ³⁺	9.4	2.2	-	-			8.2							
		Tm ³⁺	-18.7	-6.9	-	-			9.7							
		Yb ³⁺	-6.7	-2.1	-	-			8.9							
<p>Ln-C6</p> 	Ubi A28C (PDB 2MJB)	Dy ³⁺	-9.4	-5.9	-	-			10.3							
		Tb ³⁺	-14.6	-3.6	-	-			9.0							
		Tm ³⁺	11.5	4.1	-	-			9.0							
		Yb ³⁺	2.0	0.9	-	-			10.3							
<p>Ln-C7</p>	Ubi A28C (PDB 2MJB)	Dy ³⁺	26.6	6.0	-	-			4.4							
		Tb ³⁺	11.7	1.7	-	-			5.8							
		Tm ³⁺	-19.4	-7.8	-	-			6.3							
		Yb ³⁺	5.8	3.0	-	-			5.0							
		Tb ³⁺	2.2	1.0	-	-			4.9							

	GB1 Q32C (PDB 1PGA)	Tm ³⁺	-14.2	-4.4	-	-	7.3			
	HPPK S112C C80A (PDB 3QBC)	Tm ³⁺	54.5	12.5	-	-	7.3			
	Ubi A28C (PDB 2MJB)	Dy ³⁺	31.2	7.1	-	-	5.2	5.0 eq. LCT, pH 8.0, 50 mM HEPES, rt, 2 h		
		Tb ³⁺	14.3	5.1	-	-	5.0			
		Tm ³⁺	-16.3	-10.2	-	-	6.1			
		Yb ³⁺	-4.2	-1.9	-	-	5.4			
	GB1 Q32C (PDB 1PGA)	Tb ³⁺	6.1	3.1	-	-	7.4			
		Tm ³⁺	-14.9	-6.4	-	-	7.1			
	Ubi E24C A28C (PDB 2MJB)	Tm ³⁺	45.0	18.7	-	-	5.3/6.5	5.0 eq. LCT, pH 8.0, 50 mM HEPES, rt, 1 h	[145]	
		Yb ³⁺	-11.8	-4.1	-	-				
	HPPK K76C C80 (PDB 3QBC)	Tm ³⁺	47.4	19.9	-	-	6.4/7.3	3.0 eq. LCT, pH 8.0, 10 mM MgCl ₂ , 100 μM α,β-methyleneadenosine5'-triphosphate, 0 °C, 15 min		
	Ubi E24C A28C (PDB 2MJB)	Tm ³⁺	-39.9	-21.0	-38.2	-12.5	5.4/6.3	5.0 eq. LCT, pH 8.0, 50 mM HEPES, 0 °C, 1 h		
		Yb ³⁺	13.7	6.8	-	-				
	HPPK K76C C80 (PDB 3QBC)	Tm ³⁺	57.8	15.3	51.3	10.8	6.2/6.2	3.0 eq. LCT, pH 8.0, 10 mM MgCl ₂ , 100 μM α,β-methyleneadenosine5'-triphosphate, 0 °C, 15 min		
		Yb ³⁺	-9.0	-2.8	-	-				
		Ubi E24C A28C (PDB 2MJB)	Tm ³⁺	49.9	8.8	44.1	7.4	6.8/7.6	5.0 eq. LCT, pH 8.0, 50 mM HEPES, 0 °C, 3 h	[146]
			Yb ³⁺	-9.1	-4.2	-	-			
HPPK K76C C80 (PDB 3QBC)		Tm ³⁺	41.0	6.8	36.9	1.9	5.8/7.8	3.0 eq. LCT, pH 8.0, 10 mM MgCl ₂ , 100 μM α,β-methyleneadenosine5'-triphosphate, 0 °C, 15 min		
		Yb ³⁺	7.1	4.6	-	-				
Ln-T2		Tm ³⁺	35.3	4.0	29.6	9.0	5.6/6.1			

	Ubi E24C A28C (PDB 2MJJB)	Yb ³⁺	5.5	2.1	-	-	4.6/8.4	5.0 eq. LCT, pH 8.0, 50 mM HEPES, 0 °C, 3 h	
	HPPK K76C C80 (PDB 3QBC)	Tm ³⁺ Yb ³⁺	58.0 9.5	9.9 5.4	51.2 -	7.3 -		3.0 eq. LCT, pH 8.0, 10 mM MgCl ₂ , 100 μM α,β-methyleneadenosine5'-triphosphate, 0 °C, 15 min	
Ln-M7Py 	GB1 E19C (PDB 2QMT)	Tm ³⁺	33.2	12.3	-	-	7.9	4.0 eq. LCT, pH 7.6, 50 mM phosphate, 0.5 mM TCEP, 40 °C, 24 h stable under reductive conditions	[3]
	GB1 K28C (PDB 2QMT)	Tm ³⁺	-32.0	-18.2	-	-	7.7		
	GB1 E42C (PDB 2QMT)	Tm ³⁺	-33.7	-21.9	-	-	6.7		
Ln-M7FPy 	Ubi S57C (PDB 1UBI)	Tm ³⁺	17.8	9.8	-	-	6.6	6.0 eq. LCT, pH 7.0, 10 mM phosphate, 0.45 mM TCEP, rt, 16 h stable under reductive conditions	[147]
	hCA II S166C C206S (PDB 3KS3)	Tm ³⁺	-18.4	-4.4	-	-	5.4		
Ln-M7PyThiazole 	Ubi S57C (PDB 1UBI)	Tm ³⁺	35.4	11.3	25.7	3.8	5.9	4.0 eq. LCT, pH 7.0, 10 mM phosphate, 0.1 mM TCEP, rt, 15 min stable under reductive conditions	
		Dy ³⁺	41.9	26.1	-36.3	-20.0			
		Tb ³⁺	30.8	19.6	-18.6	-8.3			
	Ubi K48C (PDB 1UBI)	Tm ³⁺	47.4	8.3	46.5	12.8	7.4		
		Dy ³⁺	58.0	24.6	62.8	34.4			
		Tb ³⁺	36.0	21.8	-39.8	-24.3			
	hCA II S166C C206S (PDB 3KS3)	Dy ³⁺	44.0	24.7	42.9	22.9	7.1		
Ln-C3	Ubi G18AzF (PDB 1D3Z)	Tb ³⁺	9.0	2.3	-	-	9.1		[6]
		Tm ³⁺	-5.4	-1.7	-	-			

								10.0 eq. LCT, pH 7.5, 50 mM HEPES, 0.2 mM CuSO ₄ , 1.0 mM BTAA, 5.0 mM sodium ascorbate, rt, 16 h stable under reductive conditions	
Ln-C4 	Ubi G18A2F (PDB 1D3Z)	Tb ³⁺ Tm ³⁺	9.7 -9.7	3.4 -3.0	- -	- -	8.9		
Ln-P4T 	Ubi S57C (PDB 1UBI)	Dy ³⁺	54.0	27.0	-45.0	-21.8	6.1	6.0 eq. LCT, pH 7.0, 10 mM phosphate, 0.3 mM TCEP, rt, overnight	[150]
		Tm ³⁺	39.3	14.6	-36.5	-11.5		stable under reductive conditions	
	Ubi K48C (PDB 1UBI)	Dy ³⁺	-53.7	-23.5	-	-	6.7	stable under reductive conditions	
		Tm ³⁺	39.6	13.4	-	-		stable under reductive conditions	
	hCA II S166C C206S (PDB 3KS3)	Dy ³⁺	-46.1	-30.5	-	-	7.0	10.0 eq. LCT, pH 6.8, 10 mM phosphate, 0.3 mM TCEP, rt, overnight	
		Tm ³⁺	44.3	4.8	-	-		stable under reductive conditions	
Ln-M8-CAM-I	Ubi K6C (PDB 1D3Z)	Dy ³⁺	23	12	-	-	9.9	2.0 eq. LCT, pH 7.4, 10 mM phosphate, 140 mM NaCl, rt, 1 week	[4]
	Ubi S57C (PDB 1D3Z)	Dy ³⁺	6.8	2.4	-	-	8.0	or 10.0 eq. LCT, pH 7.4, 10 mM phosphate, 140 mM NaCl, rt, 24 h	

								stable under reductive conditions			
Ln-DO3MA-Py 	Ubi G47C (PDB 1UBI)	Tm ³⁺	22.2	13.00	-	-	8.1 (to C _A)	3.0 eq. LCT, pH 9.0-9.5, 20 mM Tris, 0.2 mM TCEP, rt, 20 h stable under reductive conditions	[151]		
	GB1 T11C (PDB 2QMT)	Tm ³⁺	21.8	13.1	-	-	8.8				
Ln-DO3MA-6MePy 	Ubi G47C (PDB 1UBI)	Dy ³⁺	-84.3	-17.2	-	-	8.9 (to C _A)				
		Tm ³⁺	65.2	31.8	-	-					
		Yb ³⁺	11.6	2.5	-	-					
	GB1 T11C (PDB 2QMT)	Dy ³⁺	51.2	20.3	-	-	6.3				
		Tm ³⁺	22.1	12.1	-	-					
CLaNP-3 	Paz E51C E54C (PDB 1PY0)	Yb ³⁺	5.9	1.9	-	-	10.0/9.0 (PRE optimized)	5.0 eq. LCT, pH 7.0, 20 mM phosphate, 4 °C, overnight	[153]		
CLaNP-5	Paz L8C K10C (PDB 1PY0)	Yb ³⁺	9.0	3.3	7.5	2.3	graphically displayed	10.0 eq. LCT, pH 7.0, 20 mM phosphate, 150 mM NaCl, 0 °C, 1 h	[24]		
		Yb ³⁺	9.0	2.3	8.4	2.0					
		Eu ³⁺	1.5	0.1	-1.4	-0.7					
		Tm ³⁺	55.3	6.9	44.4	6.9					
	Paz D100C S104C (PDB 1PY0)	Dy ³⁺	-36.2	-15.0	-34.6	-11.0					
		Yb ³⁺	9.3	3.0	5.8	1.7					
		Yb ³⁺	8.1	2.5	-	-	7.3/7.7		[157]		

	Cyt c N56C L58C (PDB 1YCC)	Tm ³⁺	50.5	6.0	-	-			
<p>CLaNP-7</p> 	Paz E51C E54C (PDB 1PY0)	Tm ³⁺	41.4	9.6	-	-	-	5.0 eq. LCT, pH 7.0, 20 mM phosphate, 150 mM NaCl, 4 °C, overnight	[155]
	Cyt c N56C L58C (PDB 1YCC)	Yb ³⁺	4.2	5.6	-	-	-		
		Yb ³⁺	6.3	8.0	-	-	graphically displayed		
<p>CLaNP-9</p> 	Paz E51C E54C (PDB 1PY0)	Yb ³⁺	7.5	2.3	-	-	-	10.0 eq. LCT, pH 8.0, 20 mM Tris, 150 mM NaCl, rt, 6 h stable under reductive conditions	[154]
		Tm ³⁺	39.5	10.6	-	-			
	T4 Lys N55C V57C (PDB 3DKE)	Yb ³⁺	6.3	2.2	-	-			
		Tm ³⁺	32.5	15.6	-	-			

3.5 Parameters in the design of LCTs

A unique feature of the structural restraints obtained by attachment of an LCT to a biomacromolecule when compared to NOESY experiments is the distance range that can be covered. The largest anisotropy parameters reported by Yang et al. for the Ln-DO3MA-6MePy LCT [151], i.e. $\Delta\chi_{ax} = -84.3 \times 10^{-32} \text{ m}^3$ and $\Delta\chi_{rh} = -17.2 \times 10^{-32} \text{ m}^3$ for a dysprosium-loaded LCT attached to an ubiquitin G47C protein construct, yield a distance limit of 207 Å assuming a threshold of significant chemical shift differences of 0.005 ppm. Based on the intrinsic tensor anisotropies currently investigated in our research group and a full transfer of the intrinsic anisotropy caused by the lanthanide to the protein by an ideal, rigid attachment of the LCT to the protein under investigation, it is not unrealistic to monitor distances over 250 Å in the future, which would greatly extend the versatility for biological applications.

The accuracy and precision in the localization of ligands within biomacromolecules that can be reached using LCTs is in the best cases comparable to the resolution of X-ray structures. As reported by Zimmermann et al., the localization of fluorinated inhibitors within human carbonic anhydrase was achieved with an accuracy of up to 0.8 Å [71]. When using LCTs for the localization of ligands within biomacromolecules, paramagnetic susceptibility tensors that intersect in a close to orthogonal fashion at the ligand position deliver the best results in terms of accuracy. To judge whether the isosurfaces of the tensors intersect in a favourable fashion, an angle score was introduced by Zimmermann et al. that allows to give a quantitative estimate about the accuracy of the ligand localization using a given set of tensors [71]. Double-armed LCTs that yield predictable tensor orientations and magnitudes [69] are particularly suited for this approach. The PCS based method opens a route to highly accurate localization of ligands within biomacromolecules even for cases that are difficult to investigate using X-ray crystallography.

For even further improvement of LCTs for biomacromolecules, three important questions in the design of LCTs emerge: i.) “What kind of ligand delivers the best framework for the synthesis of LCTs in terms of stability and inertness of the lanthanide complexes?” ii.) “How is it possible to maximize the intrinsic anisotropy of the lanthanide complexes?”, iii.) “How is it possible to transfer the full anisotropy caused by the lanthanide complex to the investigated biomacromolecule?”.

Using a rigid framework as e.g. DOTA-derived chelators, the lanthanide ion can be conveniently accommodated in the cavity and the tightly chelating ligand ensures chemical stability and inertness towards the environment of the LCT, which is a unique feature that clearly supersedes the stability of open-chain chelator-based LCTs [9].

The investigation of the intrinsic anisotropies of lanthanide complexes in combination with their structural properties is subject of current research [52, 53, 158, 159]. Furthermore, Mironov et al. reported quantitative relations between ligand atom positions and geometries and the observed anisotropies [144]. However, rational design of complexes with enhanced paramagnetic properties is cumbersome, since it is not straightforward to estimate the position of the ligand atoms in an LCT that is envisioned to be synthesized. A more convenient approach towards large intrinsic anisotropy parameters constitute the introduction of sterically demanding substituents on both the basic cyclen ring moiety as well as on the side-arms of the DOTA chelator. Rigidified DOTA-complexes provide shift ranges in one-dimensional ^1H spectra of over 1000 ppm for their Tm^{3+} and Dy^{3+} complexes, thereby indicating the large intrinsic anisotropy of the lanthanide complex [23].

In order to transfer this large intrinsic anisotropy to the protein, motional and rotational averaging of the LCT on the protein has to be suppressed or at least reduced. This was successfully achieved using two-armed LCTs [145, 146] as well as by using large substituents on the basic cyclen ring [148] or on the side-arms of the chelator [63]. In the latter two examples, the large substituents effectively decrease the rotational averaging on the proteins surface. With the current data available it is, however, not entirely clear whether substituted cyclen rings like tetramethyl-cyclen lead to improved anisotropy parameters mainly due to the suppression of the $\delta\delta\delta\delta - \lambda\lambda\lambda\lambda$ interconversion or due to the more efficient immobilization on the protein surface – probably both effects play a vital role.

The introduction of a suitable linker moiety is crucial for the successful design of an LCT. The important factors for success that have to be addressed in the design of LCTs are: i.) A short and rigid tether between the LCT and the protein should result from the tagging reaction, ii.) The chosen activator should react in fast and irreversible fashion, iii.) The resulting linking site should be reduction-stable to enable NMR studies under reductive conditions, e.g. in in-cell applications.

The LCTs presented by Lee et al. combine a very short linker with the highly reactive pyridine-2-thiol activator, that is commonly used in PCS NMR spectroscopy [25, 146]. In terms of providing a rigid linker moiety that enables fast tagging reactions resulting in a reduction-stable linkage, the pyridine-thiazolo linker equipped with a methylsulphone as leaving group reported by Müntener et al. [147] displays favourable properties: i.) Ubiquitin S57C is tagged within 15 min. in an extent of >95%, ii.) The tagging reaction as well as the NMR experiments can be performed under reductive conditions, iii.) The rigid linker moiety ensures a successful transfer of the anisotropy from the LCT to the protein resulting in large observed anisotropy parameters.

Although the limits in the design of LCTs were already pushed very far, in the near future the combination of the approaches described in this chapter and the invention of novel linker moieties enabling to suppress the remaining rotational and motional averaging could lead to a significant boost in the observed anisotropy parameters and thus to an even higher distance range covered by LCTs and a larger accuracy in the localization of ligands within biomacromolecules due to a larger magnitude of the observed paramagnetic effects.

4. Applications of lanthanide chelating tags in the structural characterization of biomacromolecules

4.1 Applications of lanthanide chelating tags on proteins I: pseudocontact shift

Since induced paramagnetic effects of LCTs, e.g. PCSs and RDCs, provide three-dimensional structural information, their use is highly valuable in the structural characterization of targets that are cumbersome to be analysed with other methods such as cryogenic electron microscopy (cryo-EM), X-ray crystallography and non-paramagnetic NMR. In a diverse range of cases, the data provided by other methods gives a valuable starting point for the structural characterization of proteins, since it allows e.g. to start from a solid state structure that can be subsequently refined to a solution state structure or e.g. flexible parts of the protein in solution can be identified.

Keizers et al. used a ClaNP-5 LCT loaded with Eu^{3+} , Gd^{3+} , Lu^{3+} and Tm^{3+} metal ions for the structural investigation of a protein complex of adrenodoxin and adrenodoxin reductase with a molecular weight of 65 kDa [160]. The observed PCSs were used as structural restraints in docking studies, in order to evaluate the binding between the two proteins in solution. The resulting structure closely matched the crystal structure of the investigated complex and the presented method holds great promise for the investigation of structures of protein complexes even at diluted conditions, since for the observation of the required PCS only ^1H - ^{15}N HSQC spectra need to be acquired.

In 2011, Cruz et al. investigated the conformational changes in the Dengue virus NS2B-NS3 protease using PCSs as structural restraints [161]. The protease was equipped with multiple LCTs on several different tagging sites and the experimental PCSs were then used for the calculation of the structure of the Dengue NS2B-NS3 protease based on the crystal structure of the related West Nile virus protease. The results show that for positively charged, low-molecular weight inhibitors, the Dengue NS2B-NS3 protease adopts a closed state resembling to the one reported for the West Nile virus protease.

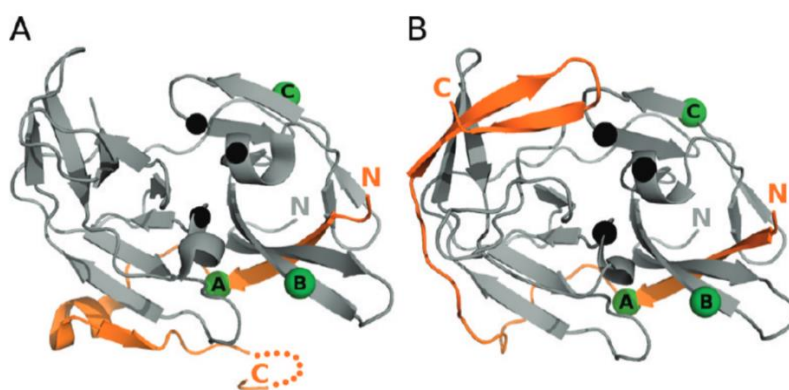


Fig. 19. Open (A) and closed (B) conformation of the C-terminal fragment of the dengue NS2B-NS3 protease investigated by Cruz et al. [161]. (Reprinted with permission from L. de la Cruz, T.H.D. Nguyen, K. Ozawa, J. Shin, B. Graham, T. Huber, G. Otting, *Binding of Low Molecular Weight Inhibitors Promotes Large Conformational Changes in the Dengue Virus NS2B-NS3 Protease: Fold Analysis by Pseudocontact Shifts*, JACS, 133 (2011) 19205-19215. Copyright 2011 American Chemical Society.)

In order to obtain a more detailed model of the catalytic domain of the matrix metalloproteinase MMP-1, Bertini et al. attached the ClaNP-5 LCT containing Lu^{3+} , Yb^{3+} , Tm^{3+} and Tb^{3+} to the H132C K136C mutant of CAT MMP-1 [162]. Importantly, Bertini et al. showed, based on the observed PCS and RDC data, that the solution structure of CAT

MMP-1 shows local rearrangements as well as significant secondary structure and domain rearrangements in solution when compared to the previously obtained crystal structure.

In 2013, it was shown by Li et al. that observation of PCS in combination with the Rosetta software package can yield 3D structures of proteins in a magic angle spinning solid-state NMR approach [163]. GB1 was tagged with 4MDPA, loaded among other paramagnetic metal ions with Yb^{3+} and Tm^{3+} . The introduction of the LCT gave rise to long-range structural restraints and the obtained structures using the observed PCSs in combination with the Rosetta database yield 3D structures similar to the X-ray structure model.

Yagi et al. attached two different LCTs (C1 and IDA-SH) to the chaperone ERp29, which is known to exhibit significant differences in the C-terminal domain between the solid state and solution state structures, and observed PCSs from four different tagging sites [164]. Based on the observed PCS data, the authors showed that the C-terminal domain of rat ERp29 in solution strongly resembles to the reported crystal structure of the C-terminal domain of human ERp29. Since only ^1H - ^{15}N HSQC experiments have to be acquired for the structure refinement, the approach could be applied to more diluted protein samples. Furthermore, the presented approach is generally applicable for protein scaffolds, in which the required tagging sites can be mutationally introduced.

In order to get structural insights into the complex of putidaredoxin and P450cam that catalyses the hydroxylation of camphor using two electron transfers, Hiruma et al. determined both the solution state structure by paramagnetic NMR as well as the crystal structure using X-ray crystallography [165]. The researchers observed 446 structural restraints upon tagging of the protein complex using the CLaNP-7 loaded with Yb^{3+} and obtained an ensemble of 10 structures with an excellent RMSD of 1.3 Å. Furthermore, the authors determined the crystal structure that strongly resembles the 10 lowest-energy structures in solution.

In 2014, Chen et al. showed that the solution structure of the C-terminal β -hairpin of NS2B in the dengue virus NS2B–NS3 protease is different from the observed conformation in the crystal structure [166]. Importantly, the C-terminal segment, which was earlier thought to be located in an open conformation, was shown by analysis of PCSs to adopt a different structure, i.e. it is closely positioned to the ligand binding site. The results provide important structural information for the design of new medications against dengue virus.

The single-stranded DNA binding protein (SSB) has two significant roles: on one hand it protects single-stranded DNA and on the other hand it recruits single-stranded DNA processing enzymes. Although the unstructured C-terminal peptide segment of SSB was shown to interact with the single-stranded DNA binding domain (OB-domain) by biochemical experiments with C-terminally truncated SSB, no evidence for competitive binding of the C-terminal peptide fragment and a single-stranded DNA molecule was found in X-ray structures. Shishmarev et al. investigated the protein using NMR spectroscopy and showed, based on NOESY data as well as PCSs induced by a lanthanide-loaded IDA-SH tag [128], that the C-terminal fragment indeed binds to the OB-domain, suggesting direct competition between the C-terminal peptide fragment and single-stranded DNA in binding to the OB-domain [167].

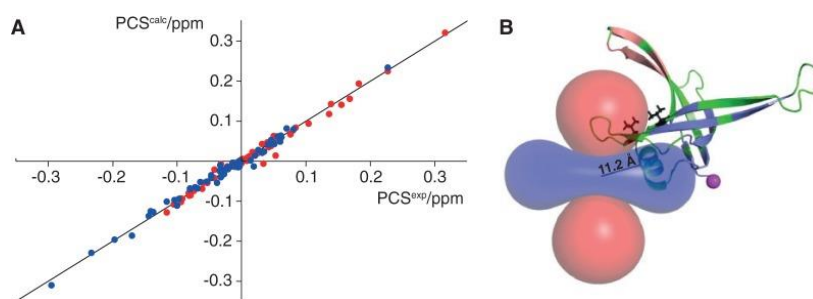


Fig. 20. PCS correlation plot measured on SSB (A, red: Tb^{3+} , blue: Tm^{3+}). Orientation of the isosurfaces of the fitted anisotropy parameters shown together with the structure of SSB (B) [167].

Inspired by the results of Chen et al. [166], de la Cruz et al. tagged a two-component dengue virus NS2B-NS3 protease with two different paramagnetic metal tags [168]. In order to validate the earlier results from X-ray crystallography in solution, i.e. the arrangement of NS3 protease and NS2B, NS3 protease was equipped with paramagnetic LCTs and PCSs were analysed to yield a magnetic susceptibility tensor. The results clearly show that the reported structures depict the interaction between NS2B and NS3 protease in a correct fashion. Furthermore, the authors showed that NMR signals, that were believed to originate from the open state conformation, are artefacts from protein degradation.

Camacho-Zarco et al. presented in 2014 the use of a reporter protein tagged with a Tm^{3+} containing ClaNP-5 LCT [169]. The paramagnetically labelled PDZ mutants can then be fused to another protein of interest containing a C-terminal TGWETWV peptide sequence, giving rise to observation of paramagnetic effects on the unlabelled protein. While the sparse nuclear Overhauser effect (NOE) signal sets failed in the 3D structure determination of the 7TM α -helical microbial receptor pSRII, Crick et al. showed that introduction of a paramagnetic metal tag yields PCS restraints that allow to determine the correct three-dimensional fold of the membrane protein [170]. Since the limitation of the NOE signal sets displays an often encountered problem in membrane proteins incorporating α -helical structures, the authors suggest that the acquisition and analysis of PCSs is a general method to overcome this issue.

Analysis of PCSs by Brewer et al. solved a long-standing biological riddle and led to structural characterization of the neurotransmitter release triggered by Syt1, a Ca^{2+} sensor that together with the SNARE complex (synaptobrevin, syntaxin-1 and SNAP-25) is responsible for rapid neurotransmitter release [12]. The NMR data suggest a dynamic binding mode in which basic residues of the Syt1 C2B-domain β -sandwich bind to a polyacidic part of the SNARE complex. Furthermore, Pan et al. demonstrated that it is possible to assign the methyl groups of the neuronal SNARE complex, i.e. the leucine, isoleucine, methionine and valine methyl groups, in a sequence-specific manner based solely on measurement of lanthanide-induced PCSs [171].

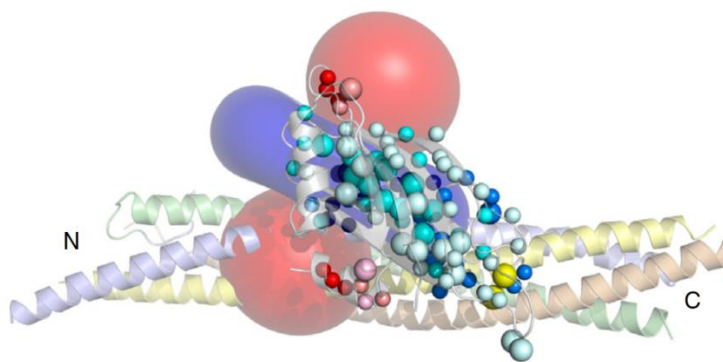


Fig. 21. Models of Syt C2B bound to the SNARE complex matching the Syt C2B PCSs induced by an LCT bound to residue 166 [12].

In 2015, Saio et al. reported that the detailed conformational characterization of the multidomain protein MurD with a molecular weight of 47 kDa was enabled by PCS analysis [172]. In order to detect sizeable PCSs, ClaNP-5 loaded with Yb^{3+} and Tm^{3+} was introduced into the protein by ligation to the cysteine residues of the mutant MurD E260C K262C. Upon analysis of the obtained structural restraints, the researchers found evidence for a novel semi-closed conformation that can exhibit different binding affinities compared to the open or closed conformation of the protein. The results are of high importance in the design of novel antibiotics in fragment based drug discovery, since, based on data from X-ray crystallography, MurD was previously thought to adopt only an open or closed conformation.

Pilla et al. demonstrated the use of LCTs for the generation of three-dimensional structure models of proteins larger than 200 residues relying solely on PCSs induced from 4 different tagging sites [105]. The new algorithm that relies on the Rosetta approach yields fragment libraries that are then refined upon convergence. The generation of three-dimensional models of protein structures by PCSs and a Rosetta fragment library as only input holds great promise for the structural investigation of proteins that are difficult to characterize by X-ray crystallography, cryo-EM and NOE-based methods.

Chen et al. found that for the bacteriophage T4 lysozyme (T4-L), which shows two subdomains that are arranged in two different conformations in its crystal structure, only one set of signals is present in NMR experiments, pointing out conformational homogeneity [173]. In order to validate the results, a 4MDPA-derived LCT was attached to the protein. The obtained PCSs and RDCs strongly suggest that in absence of the substrate, the T4 lysozyme adapts in contrast to the X-ray structure a more open conformation.

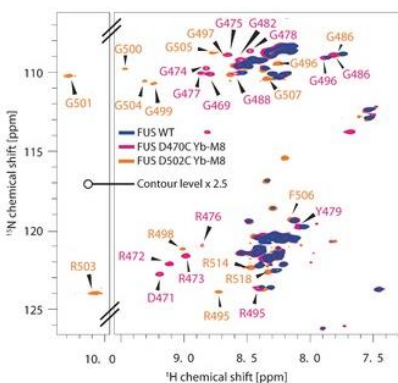


Fig. 22. Increasing the chemical shift dispersion of IDPs using an Yb-DOTA-M8 LCT as demonstrated by Göbl et al (spectra acquired at 900 MHz proton frequency) [73].

Intrinsically disordered proteins are challenging subjects for the structural characterization by NMR due to the significant overlap of resonances and motions in the intermediate exchange regime, which can lead to signal broadening. In order to overcome the severe overlap of signals, Göbl et al. demonstrated in 2016 successfully the use of an Yb³⁺ loaded DOTA-M8 LCT and were thereby able to increase the chemical shift dispersion, greatly facilitating analysis of the otherwise overlapping signals [73].

Although low-populated excited states can yield valuable information about a protein's function, it can be a challenging task to monitor these low abundance states and derive structural models from the data set. Ma et al. demonstrated that PCSs acquired via chemical exchange saturation transfer (CEST) provide a useful tool for both the acquisition of the signals of the otherwise invisible state as well as the structural modelling of the protein conformation [174]. The authors applied their approach in the characterization of the bound state of 3 mol% Ark1p peptides to the Abp1p SH3domain and the detection of transitional conformations of the FF domain.

In order to challenge the potential of LCTs for the determination of three-dimensional structures of proteins, Pearce et al. demonstrated that with only two different tagging sites and the application of one lanthanide ion, namely Tb³⁺, the three-dimensional structure of the model protein ubiquitin can be investigated [26]. The authors of the study found that a previously published ensemble structure of ubiquitin fits better to the paramagnetic restraints than structures with only one conformation, a result that points out that PCSs can be used to detect very subtle changes in the conformation of biomacromolecules.

Lescanne et al. reported the assignment of methyl groups within the 25 kDa N-terminal domain of Hsp90 achieved by the analysis of PCSs induced by Yb³⁺-loaded CLaNP-5 [24] from two or three different tagging sites using a novel version of the program PARAssign [175]. 60% of the 76 methyl groups were assigned in reliable manner, while for the majority of the remaining 40% only three or less potential assignments were left. In order to provide a complete assignment of the methyl groups, the data obtained from PCS analysis can then be combined with additional restraints, e.g. from NOE experiments or mutagenesis studies. In a further study, the authors demonstrated that the PCS of the methyl groups can then be used to monitor the reorientation of methyl groups upon binding of a 4-(2-fluorophenyl)-2-pyrimidinamine ligand [176]. By application of their method, the authors found displacements of the

methyl groups in the binding site of 1-3 Å and showed thereby that conformational changes of the side chains within a protein can be monitored using solely two-dimensional ^1H - ^{13}C HSQC spectra.

4.2 Applications of lanthanide chelating tags on proteins II: residual dipolar couplings

In 2007, Kamen et al. reported the use of an LCT for the characterization of membrane proteins [177]. Since the NMR signals are significantly broadened due to the large size of the protein and the use of a phospholipid bicelle or a solubilizing detergent micelle, deuteration of the investigated protein is required and hence NOE-derived structures rely only on NOE signals of backbone amides. While polyacrylamide gels cause fewer observed peaks by introducing additional line-broadening, EDTA-derived LCTs offer a robust method for partial alignment of the investigated membrane proteins and a successful measurement of ^1H - ^{15}N , ^{13}C - ^{15}N and $^{13}\text{C}'$ - $^{13}\text{C}^\alpha$ RDCs in 3D-experiments was demonstrated.

Xu et al. presented the use of an LCT attached to yeast cytochrome c in the study of a cytochrome c-adrenodoxin complex [178]. The used tag, i.e. CLaNP-5, did not interfere with the binding and showed the highly dynamical nature of the formed protein complex. While the detected RDCs for cytochrome c led to successful determination of a magnetic susceptibility tensor, the observed RDCs on adrenodoxin were significantly smaller than expected, indicating multiple conformation and dynamical averaging.

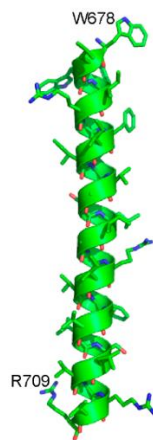


Fig. 23. Model of the transmembrane domain of the HIV-1 viral envelope protein gp41 modelled using RDCs by Chiliveri et al. [179]. (Reprinted with permission from S.C. Chiliveri, J.M. Louis, R. Ghirlando, J.L. Baber, A. Bax, *Tilted, Uninterrupted, Monomeric HIV-1 gp41 Transmembrane Helix from Residual Dipolar Couplings*, JACS, 140 (2018) 34-37. Copyright 2018 American Chemical Society.)

Although experimental evidence from cryo-EM and X-ray crystallography suggests a homotrimeric, C_3 symmetric structure of the transmembrane domain of the HIV-1 viral envelope protein gp41, Chiliveri et al. showed by the use of RDCs observed under three different alignment conditions (neutral and positively charged acrylamide gels as well

as paramagnetic alignment using a thulium containing LCT) and DEER EPR, that the transmembrane domain is monomeric and exhibits a highly ordered element of 32 residues [179].

In 2019, Barnes et al. showed by introduction of a thulium loaded DOTA-derived LCT that the 68-residue long medial-tail domain of myosin-VI, which corresponds to a highly charged single α -helical domain (SAH domain), adopts in solution the structure of a highly ordered α -helix [180]. Residues more distant from the attachment site show a decrease in alignment and these values were used for an exact calculation of the persistence length (defined as the distance in a polymer upon which correlations are lost) of the SAH helix. The result of 224 Å assists the hypothesis that the investigated SAH helix is sufficiently stiff to be a part of the lever arm of myosin-VI.

4.3 Applications of lanthanide chelating tags on proteins III: protein-ligand complexes

Although the localization of ligands within proteins is highly important for any studies relating to drug molecules and their associated mechanism, still by today it is not straightforward to localize ligands on proteins that are not suitable for crystallization in complex with the protein and readily subjectable to X-ray crystallography. Ligand positions can be identified using NMR techniques as CSP or NOESY experiments. While CSP can be misleading or yield ambiguous results due to conformational changes occurring in the protein upon ligand binding or unexpected intermolecular interactions, NOESY spectra and their analysis require in most cases a complete NMR assignment of the accordingly labelled protein of interest. PCS analysis can provide the needed 3D information in a GPS-like approach from simple one-dimensional NMR experiments.

In 2011, Saio et al. showed that localization of ligands on proteins tagged with an LCT is possible by the structural restraints obtained by PCS [70]. Upon tagging of the protein of interest and determination of the magnetic susceptibility tensors associated with the different tagging sites, the observed PCSs on the ligand atoms were translated in distance restraints. In combination with a docking simulation, Saio et al. successfully localized a high-affinity inhibitor within Grb2 SH2 over a distance of ~ 10 Å.

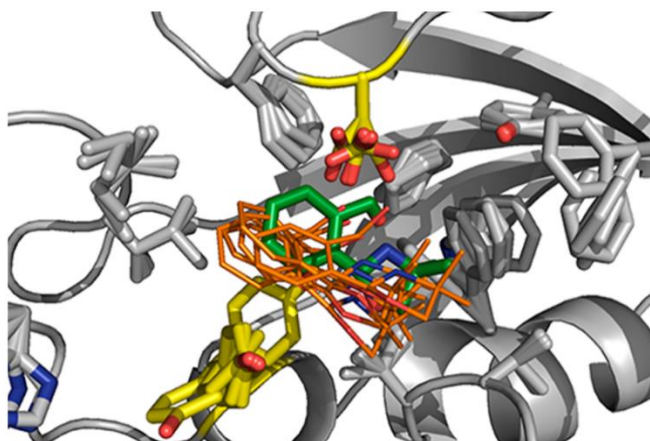


Fig. 24. Localization of a ligand on a protein by PCS reported by Guan et al. (green: averaged NOE structure, orange: best five obtained structure by PCS-fitting) [69]. (Reprinted with permission from J.-Y. Guan, P.H.J. Keizers, W.-M. Liu, F. Löhner, S.P. Skinner,

In contrast to strongly binding ligands, weakly binding ligands can be applied in an excess and hence the needed signal discrimination between the protein and the applied ligand is more conveniently achieved than in the case of strong binders. From the fraction of bound ligand, the PCSs can be extracted. Guan et al. demonstrated this methodology for the localization of weakly binding ligands on a protein using two-point anchored LCTs [69]. Two-armed lanthanide probes induce, due to their high immobilization on the protein scaffold, large PCSs and provide the additional advantage that the tensors of the anisotropic part of the magnetic susceptibility can be predicted computationally in order to yield a low-resolution model of the ligand localization, which can be e.g. used to discriminate between two binding sites within a protein. The study performed by Guan et al. relied only on one-dimensional ^1H spectra and the observed PCSs were then used as docking restraints.

In 2015, Brath et al. presented an approach for the localization of ligands within proteins that relies on attachment of the LCT to the ligand [67]. The applied methodology presents several advantages: i) the solubility in aqueous solution of hydrophobic ligands can be enhanced and study of their interactions with proteins is facilitated, ii) engineering of protein cysteine mutants is omitted, iii) multiple binding sites could easily be detected, iv) if the ligand is sufficiently large, it can be ensured that no structural changes of the protein or the protein-ligand binding complex occur, and hence the interactions of the ligand with the native protein can be detected. However, chemical modification/tagging of the ligand renders the presented method less suitable for screening of a large number of ligands within a protein, in contrast to an approach that relies on tagging of the protein and is able to screen a large number of ligands with one tagged protein construct.

Chen et al. reported in 2016 an interesting approach to overcome the problem of discrimination between signals of a strongly binding ligand and the protein [181]. A *tert*-butyl group was used as a reporter group on the used ligand, since the *tert*-butyl group is represented in spectra with an extremely narrow signal and a specific NMR shift, i.e. the signal lies in the region of 1.0-1.5 ppm. Furthermore, the authors of the published study found that the *tert*-butyl group gives rise to significant signals in NOESY spectra, even when protein NOESY correlations are hardly detectable. The used approach constitutes a very convenient approach for the investigation of complexes of strongly binding, *tert*-butyl group containing ligands and a protein of interest.

Gao et al. demonstrated in 2017 the use of a ^{19}F -PCS CEST approach in order to characterize protein-ligand interactions in the NMR intermediate exchange regime [66]. By scanning through the ^{19}F saturation frequencies and detection of the free ligand, intensity dips can be measured at the frequency of the resonance of the bound state. By use of their method, the researchers localized mono- and di-fluorinated inhibitors of the BRM bromodomain by PCS combined with docking approaches and confirmed the results by analysis of the solved crystal structure. Importantly, upon initial determination of the magnetic susceptibility tensors, a large number of ligands can be screened without further need for protein engineering.

In 2018, Xu et al. presented a method suitable for the localization of ligands exhibiting micromolar affinities [182]. Since ligands with micromolar affinities can fall in the limit of NMR intermediate exchange and ^1H CEST usually suffers from NOE interferences, the authors acquired the required spectra using relaxation dispersion. The researchers successfully observed the desired PCS and used them as docking restraints for the generation of structures of the protein-ligand complexes.

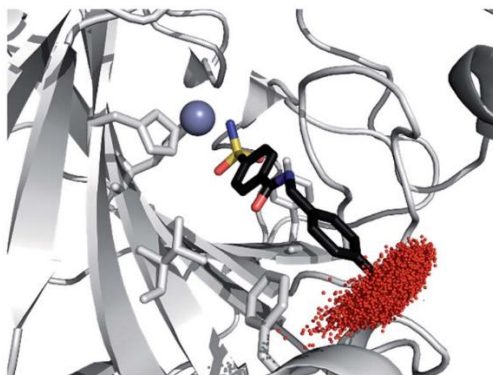


Fig. 25. Point cloud of the Monte–Carlo fluorine position calculation for FM-520 within hCA-II (light blue sticks (hidden by red point cloud): fluorine atoms, blue sphere: Zn^{2+} ion) [71]. (K. Zimmermann, D. Joss, T. Müntener, E.S. Nogueira, M. Schäfer, L. Knörr, F.W. Monnard, D. Häussinger, Localization of ligands within human carbonic anhydrase II using ^{19}F pseudocontact shift analysis, *Chem. Sci.*, 10 (2019) 5064–5072. - Published by The Royal Society of Chemistry)

Zimmermann et al. reported that ^{19}F PCS NMR can be used as a method to determine fluorine positions of high affinity ligands located within hCA II with an accuracy of up to 0.8 Å [71]. Upon determination of 4–5 independent magnetic susceptibility tensors induced by LCTs attached to different cysteine mutants, least-squares minimization and a Monte-Carlo approach allowed to localize the ligands within the protein. The unprecedented distance range of 22–38 Å holds great promise for investigation of protein-ligand complexes significantly larger than 30 kDa.

4.4 Applications of lanthanide chelating tags on carbohydrates

Since in sugar molecules often a severe signal overlap of virtually identical domains is encountered and often renders conventional J-coupling and NOE-based NMR techniques difficult to apply, introduction of LCTs and subsequent analysis of PCSs and RDCs can increase the shift dispersion and yield structural restraints that result in three-dimensional models and complement modelling studies.

In 2011, the first rigid LCT, which was covalently attached to a carbohydrate molecule, was reported by Mallagaray et al. [183]. In their study, the researchers demonstrated the successful transfer of the methodology, which was originally developed for proteins, by covalent attachment of an EDTA-based LCT via a rigid linker to 1- β -aminochitobiose and subsequent analysis of the observed PCSs.

Erdélyi et al. extended the methodology by attaching an EDTA-based chelator loaded with six different lanthanides to a lactose molecule, enabling an analysis of the conformational space of lactose [184]. The obtained restraints from

RDCs and PCSs agree with MD simulations that predicted that lactose is present in several different conformations of the glycosidic bond in aqueous environment and exclude the possibility that lactose is present in solution in a single and rigid conformation.

The methodology was then extended by Yamamoto et al. to a typical core structure of N-linked oligosaccharides (N,N'-diacetylchitobiose) [185]. Upon attachment of the EDTA-based LCT to the sugar molecule via a peptide coupling reaction employing HATU as coupling reagent, significant PCSs in ^1H - ^{13}C HSQC spectra were observed. The experimentally observed PCSs were used for a comparison with molecular-dynamics simulations and both approaches suggest a conformational rigidity of the glycosidic linkage in the investigated disaccharide and thereby of the common innermost part of N-linked oligosaccharides. The authors suggested that the presented method could be extended and interactions of N-linked oligosaccharides with lectins could be monitored using strongly paramagnetic lanthanides as Dy^{3+} and Tb^{3+} .

The attachment of a rigid LCT bound to a branched tetra-saccharide of biological importance, i.e. the branched tetra-saccharide of GM2 ganglioside was presented by Zhang et al. [186]. The combination of molecular dynamic (MD) simulations and PCS analysis led to a model of the conformational space of GM2 tetra-saccharide. The comparison with its non-branched derivative, the GM3 trisaccharide, suggests that the additional GalNAc branch restricts the conformational freedom of the Neu5Ac-Gal glycosidic linkage in the GM2 tetra-saccharide.

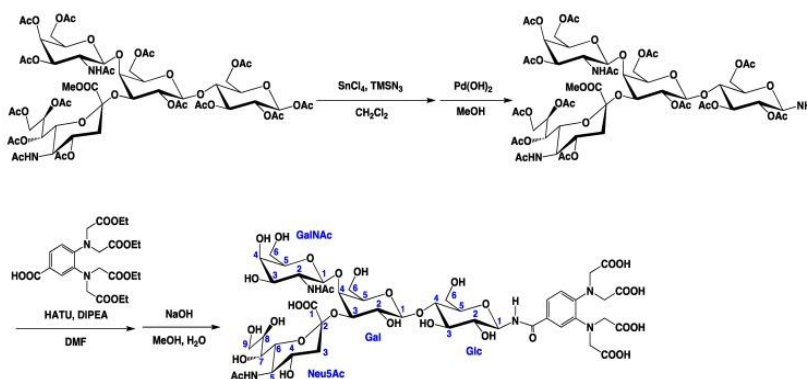


Fig. 26. Preparation of the branched tetra-saccharide covalently bound to a chelator reported in Zhang et al. [186].

Intricate N-glycans are important recognition molecules related to a diverse range of physiological pathways and diseases [187, 188]. The structural investigation of the multiple and almost identical side-chains is challenging by standard methods as non-paramagnetic NMR and X-ray crystallography. By introduction of an LCT, Canales et al. could distinguish the two arms of a symmetrical, branched N-glycan, which shows upon attachment of an LCT individual resonances for the two arms that exhibit otherwise isochronous signals, and thereby confirm the major conformation and rotamer in solution [189]. Furthermore, the approach allowed to characterize the binding epitopes of the symmetrical glycan and showed that both of the arms of the N-glycan are involved in the interaction with human galectin-3. In a further study, Canales et al. showed that it is not only possible to determine carbohydrate conformations upon attachment of an LCT to a sugar molecule, but also to monitor the interaction site in a protein

and characterize the topological changes upon binding of a model saccharide equipped with an LCT [190]. Besides the aspect that the presented method could also easily reveal secondary binding sites, the chemical modification of the sugar molecule with an LCT omits the need for protein engineering. Canales et al. demonstrated, that paramagnetic NMR allows to increase the chemical shift dispersion of the virtually identical side-chains and the identification of the four N-acetyl lactosamine antennae recognized by two N-acetyllactosamine-binding lectins, specifically *Datura stramonium* seed lectin and *Ricinus communis* agglutinin [74]. After successful separation of the isochronous signals caused by PCS and structural characterization, addition of the lectin allowed the assignment of the major lectin-carbohydrate interactions by observation of the line broadening on the different branches of the tetra-antennary N-glycans.

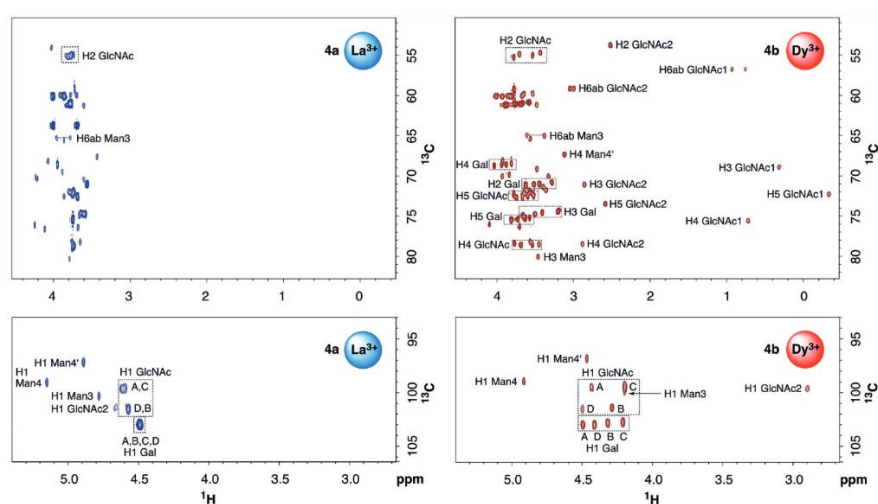


Fig. 27. Increasing the chemical shift dispersion in ^1H - ^{13}C HSQC spectra for the isochronous signals of the four N-acetyl lactosamine antennae as demonstrated by Canales et al (acquired at 600 MHz proton frequency) [74].

Mallagaray et al. ligated an LCT via a copper-catalysed azide-alkyne click reaction to *L*-fucose [191]. Interestingly, the paramagnetic carbohydrate-LCT conjugate can then be used to aid the backbone assignment of the protruding domains of a human norovirus strain that constitutes a challenging target to be investigated by NMR spectroscopy due to the high molecular weight of 70 kDa. The researchers found that the carbohydrate-LCT conjugate at low concentrations exclusively binds to the primary binding pocket. However, since peak broadening of multiple resonances was observed and compromised the use for the backbone assignment of this challenging target, synthesis of ligands with higher binding affinity or the use of covalently bound tags on the biomacromolecule scaffold are highly desirable.

4.5 Applications of lanthanide chelating tags on nucleic acids

The structure of nucleic acids in solution is essential to their function and stability [192]. In order to investigate the structural properties, NMR offers a non-destructive approach to measure the solution state structure of DNA and RNA that not only allows to measure these biomacromolecules in physiological conditions, but also to tune the buffer salt concentration, pH value and additives in virtually any possibility [193, 194]. The wide range of LCTs that was developed for the investigation of structure and dynamics of proteins includes a large variety of different attachment systems (linker moieties forming disulphide bonds [9, 24], bio-orthogonal linkers for ligation to unnatural amino acids [6], as well as reductively stable linkers for in-cell NMR [2-4]). Extension of this toolbox by new linkage systems or exploitation of the currently developed LCTs in order to study the structural properties of DNA and RNA is an important goal for future developments in the research field of paramagnetic NMR. By paramagnetic labelling of nucleic acids and derivatives, not only structural restraints for the assessment of computational studies, refinement of structures obtained by X-ray crystallography and *de novo*-derived structures can be obtained, also ligand binding positions and characterization of ligand-nucleic acid complexes can be achieved.

In 1997, Gochin et al. presented the first application of paramagnetic NMR on nucleic acids, when they investigated the formation of a ternary drug-DNA-divalent cation complex composed of a nucleic acid strand, two chromomycin-A3 molecules and a Co^{2+} (paramagnetic) or Zn^{2+} ion (diamagnetic) [195]. Subsequently, the signals were assigned and the magnetic susceptibility tensor parameters were derived [196]. Furthermore, the structural restraints gained by analysis of the PCSs were used in energy minimization and MD calculations and shown to outperform structural restraints obtained from NOESY [29].

Further initial examples of the application of paramagnetic NMR applied to nucleic acids and their derivatives include the Cys-Ph-TAHA LCT, that was covalently attached to a ternary complex of the protein LacR, the O2 operator and IPTG in 2011 by Peters et al. [14] and the structural investigation of Subp2-R3H in complex with deoxyguanosine 5'-monophosphate (dGMP) as structural analogue to the 5' end of single-stranded DNA in 2012 by Jaudzems et al. [197]. In the latter example, paramagnetic NMR allowed an insight into structural changes of the protein upon binding of the DNA mimic. Strickland et al. presented in 2019 that attachment of DOTA-M8-SSPy LCTs to a spliceosomal U1A RNA-binding domain allows for the structural characterization of a 232-nt RNA derived from the HIV-1 Rev response element [198]. Furthermore, the method was validated using a 36-nt RNA with known structural properties. Based on the fact that the seven nucleotide long U1A binding sites can be readily inserted in hairpin or loop structures, the authors suggest that the presented method is generally applicable for the atomic-level study of large RNAs.

The mentioned examples differ in the mode of the LCT-biomacromolecule linkage, i.e. in the first three examples affinity binding of the paramagnetic metal centre is exploited, while Peters et al., Jaudzems et al. and Strickland et al. attached the LCT covalently to the protein interacting with nucleic acid or to the nucleic acid derivative.

In order to prevent motional averaging of the paramagnetic effects observed in affinity binding, the covalent attachment of the LCT incorporating the paramagnetic metal centre directly to the nucleic acids constitutes a highly desirable approach. The covalent labelling of nucleic acids was demonstrated for an EPR application by Song et al. [199] and Goldfarb [200] via a copper catalysed azide-alkyne click reaction of a DNA strand incorporating an azide

functionality and an LCT bearing an alkyne functionality as well as by Wu et al. [30] for the analysis of PCSs via modification of a phosphorothioate group in DNA with an LCT containing an electrophilic bromoacetamide functionality for alkylation.

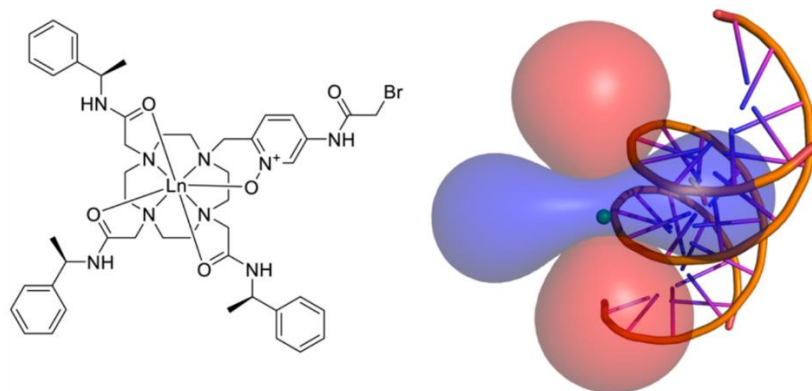


Fig. 28. Structure of the bromoacetamide LCT (left) employed by Wu et al. and isosurfaces of the obtained anisotropy parameters after successful ligation of the LCT to a phosphorothioate modified nucleophilic site in DNA (right) [30]. (Reprinted with permission from Z. Wu, M.D. Lee, T.J. Carruthers, M. Szabo, M.L. Dennis, J.D. Swarbrick, B. Graham, G. Otting, *New Lanthanide Tag for the Generation of Pseudocontact Shifts in DNA by Site-Specific Ligation to a Phosphorothioate Group*, *Bioconjugate Chem.*, 28 (2017) 1741-1748. Copyright 2017 American Chemical Society.)

Insights into the solution structure and dynamics of RNA and DNA molecules obtained by paramagnetic NMR are highly desirable not only for fundamental research, but can also give significant inputs for characterization of binding sites of RNA and DNA targeting molecules in future drug discovery. The information obtained by paramagnetic NMR can refine X-ray structures as well as complement or support modelling studies of the structure of nucleic acids. Since the application of LCTs for proteins is well established and documented, it is highly promising to transfer this knowledge to applications involving nucleic acids in the next decade. Crucial factors for a successful transfer of these tools from proteins to nucleic acids will be the development of synthetically accessible LCTs bearing linkers suitable for a covalent attachment of the LCT to the nucleic acid molecule. The structural insights gained by the exploitation of the paramagnetic effects of LCTs can then provide important data for fundamental research as well as catalyse the development of DNA and RNA targeting treatments in drug discovery.

5. Conclusion & Outlook

To conclude, over the last two decades an overwhelming variety of LCTs was developed, each with its own characteristics and fields of application. In contrast to the initial developments in the field that included LCTs hampered in application by displaying multiple signal sets and strongly averaged paramagnetic effects, the latest generations of LCTs available by today comprise LCTs that bear reduction-stable linkers and display only one single set of peaks. The recently published tags show very large anisotropy parameters ($\Delta\chi_{ax} > 40 \times 10^{-32} \text{ m}^3$) and, thus, provide extremely large PCSs and RDCs, giving rise to precise structural restraints and convenient analysis of the acquired spectra. Assuming a detection limit of 0.005 ppm for chemical shift differences, current LCTs potentially provide structural restraints over a distance exceeding 200 Å. In combination with a sufficient number of suitable protein constructs to achieve the prediction of a spin position with an accuracy of up to 0.8 Å, this constitutes PCS as a very powerful tool for future applications. The possible spectrum of applications of LCTs is extremely broad and includes the investigation of protein structures in solution, protein structures in living cells, the structural characterization of flexible domains and conformations, the identification and investigation of ligand-binding sites, the structure elucidation of protein-protein complexes, as well as applications on other biomacromolecules, such as DNA, RNA and oligosaccharides. Most important, LCTs can aid to refine structures of biomacromolecules in solution under physiological conditions and as demonstrated in various projects, can lead to completely new and surprising findings that cannot be revealed by X-ray structures or NOE studies. By giving a more detailed mechanistic insight into enzyme functions and alignment of protein subunits in solution, lanthanide chelating tags contribute significantly to the current knowledge of the structure of biomacromolecules and can lead to breakthroughs in current drug development research. In near future, further development of LCTs with suitable linkers for the investigation of DNA and RNA are highly desirable, as well as the application of the existing LCTs for the solution of long-standing problems in structural biology, especially those associated with disease mechanisms and biomedical applications. Open questions on the design of LCTs, however, remain: e.g. there is no systematic study on the influence of donor set atoms on the anisotropy parameters; the motional degrees of freedom for a given tag are still not entirely understood and many alternative chelators to DOTA as basic scaffold have not yet been thoroughly tested for PCS – so there is still work to be done.

6. Acknowledgement

The Chemistry Department of the University of Basel is acknowledged for financial support. T. Müntener and R. Vogel are acknowledged for helpful discussions.

7. Glossary

3MDPA: 3-mercapto-2,6-pyridinedicarboxylic acid

4MDPA: 4-mercapto-2,6-pyridinedicarboxylic acid

4MMDPA: 4-mercaptomethyl-dipicolinic acid

4MTDA: 4'-mercapto-2,2': 6',2''-terpyridine-6,6''-dicarboxylic acid

4PS-DPA: (4-(phenylsulphonyl)-pyridine-2,6-dicarboxylic acid

4PS-PyMTA: 4-phenylsulphonyl-(pyridine-2,6-diyl)bismethylenenitrilo tetrakis(acetic acid)

4VPyMTA: 4-vinyl(pyridine-2,6-diyl)bismethylenenitrilo tetrakis(acetic acid)

7TM: seven-helix transmembrane

ArgN: N-terminal domain of the arginine repressor from *E. coli*

AzF: *p*-azido-*L*-phenylalanine

BRM: Brahma (*Drosophila* protein)

BTAA: 2-(4-((bis((1-(*tert*-butyl)-1H-1,2,3-triazol-4-yl)methyl)amino)methyl)-1H-1,2,3-triazol-1-yl)acetic acid

CEST: chemical exchange saturation transfer

CLaNP: caged lanthanide NMR probe

Cryo-EM: cryogenic electron microscopy

CS: contact shift

CSP: chemical shift perturbation

Cyt c: cytochrome c

DEER: double electron-electron resonance

DFT: density functional theory

dGMP: deoxyguanosine 5'-monophosphate

DNA: deoxyribonucleic acid

DOTA: 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid

DPA: dipicolinic acid

DTPA: diethylenetriaminepentaacetic acid

DTT: dithiothreitol

DTTA-C3-yne: diethylene-triamine-tetraacetate propyl-1-yne

DTTA-C4-yne: diethylene-triamine-tetraacetate butyl-1-yne

ECAD12: first two domains of epithelial cadherin

EDTA: ethylenediaminetetraacetic acid

EPR: electron paramagnetic resonance

ERp29: endoplasmic reticulum protein 29

FM-520: 4-sulphamoyl-N-(4-(trifluoromethyl)benzyl) benzamide

GB1: B1 domain of protein G

GM2 and GM3: second and third discovered monosialic ganglioside

GPS: global positioning system

Grb2 SH2: SH2 domain of growth factor receptor-bound protein 2

HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate

hCA II: human carbonic anhydrase II

HeLa cell: immortal cell line derived from cervical cancer taken from Henriette Lacks

HEPES: 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulphonic acid

HPLC: high-performance liquid chromatography

HPPK: ATP:2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine 6'-diphosphotransferase

Hsp90: heat shock protein 90

HSQC: heteronuclear single quantum coherence

HSQC IPAP: heteronuclear single quantum coherence in-phase/anti-phase

IDA: iminodiacetic acid

IPTG: isopropyl β -D-1-thiogalactopyranoside

LacR: lac repressor

LBP: lanthanide binding peptide

LCT: lanthanide chelating tag

MD: molecular dynamics

MMP-1: matrix metalloproteinase 1

MRI: magnetic resonance imaging

MurD: UDP-N-acetylmuramoylalanine--D-glutamate ligase

NMR: nuclear magnetic resonance

NOE: nuclear Overhauser effect

NOESY: nuclear Overhauser effect spectroscopy

NTA: nitrilotriacetic acid

OB-domain: DNA binding domain of single-stranded DNA binding protein

Paz: domain of the proteins Piwi, Argonaut and Zwill

PCS: pseudocontact shift

PDZ domain: domain found in post synaptic density protein, Drosophila disc large tumour suppressor and zonula occludens-1 protein

PDZ protein: family of proteins containing the PDZ domain

PRE: paramagnetic relaxation enhancement

pSRII: phototaxis receptor sensory rhodopsin II

RDC: residual dipolar coupling
RMSD: root mean square deviation
RNA: ribonucleic acid
SAH domain: single α -helical domain
SAP: square antiprism
SH2 domain: Src homology 2 domain
SNAP-25: synaptosomal nerve-associated protein 25
SNARE: SNAP receptor
Src: proto-oncogene tyrosine-protein kinase (Src derived from sarcoma)
SrtA: sortase A
SSB: single-stranded DNA binding protein
Syt1: synaptotagmin 1
S μ bp2-R3H: R3H domain of S μ bp2
T4 Lys: T4 lysozyme
TAHA: triaminohexaacetate
TCEP: 3,3',3''-phosphanetriyltripropanoic acid
Tris: 2-amino-2-(hydroxymethyl)propane-1,3-diol
TSAP: twisted square antiprism
U1A: small nuclear ribonucleoprotein polypeptide A
UBA1: ubiquitin-like modifier activating enzyme 1
Ubi: ubiquitin

8. References

- [1] C. Nitsche, G. Otting, Pseudocontact shifts in biomolecular NMR using paramagnetic metal tags, *Prog. Nucl. Magn. Reson. Spectrosc.*, 98-99 (2017) 20-49.
- [2] B.-B. Pan, F. Yang, Y. Ye, Q. Wu, C. Li, T. Huber, X.-C. Su, 3D structure determination of a protein in living cells using paramagnetic NMR spectroscopy, *Chem. Commun.*, 52 (2016) 10237-10240.
- [3] T. Müntener, D. Häussinger, P. Selenko, F.-X. Theillet, In-Cell Protein Structures from 2D NMR Experiments, *J. Phys. Chem. Lett.*, 7 (2016) 2821-2825.
- [4] Y. Hikone, G. Hirai, M. Mishima, K. Inomata, T. Ikeya, S. Arai, M. Shirakawa, M. Sodeoka, Y. Ito, A new carbamidomethyl-linked lanthanoid chelating tag for PCS NMR spectroscopy of proteins in living HeLa cells, *J. Biomol. NMR*, 66 (2016) 99-110.
- [5] W.-M. Liu, M. Overhand, M. Ubbink, The application of paramagnetic lanthanoid ions in NMR spectroscopy on proteins, *Coord. Chem. Rev.*, 273-274 (2014) 2-12.
- [6] C.T. Loh, K. Ozawa, K.L. Tuck, N. Barlow, T. Huber, G. Otting, B. Graham, Lanthanide Tags for Site-Specific Ligation to an Unnatural Amino Acid and Generation of Pseudocontact Shifts in Proteins, *Bioconjugate Chem.*, 24 (2013) 260-268.
- [7] G. Otting, Protein NMR using Paramagnetic Ions, *Annu. Rev. Biophys.*, 39 (2010) 387-405.
- [8] X.-C. Su, H. Liang, K.V. Loscha, G. Otting, [Ln(DPA)₃]³⁻ Is a Convenient Paramagnetic Shift Reagent for Protein NMR Studies, *JACS*, 131 (2009) 10352-10353.
- [9] D. Häussinger, J.-R. Huang, S. Grzesiek, DOTA-M8: An Extremely Rigid, High-Affinity Lanthanide Chelating Tag for PCS NMR Spectroscopy, *JACS*, 131 (2009) 14761-14767.
- [10] X.-C. Su, B. Man, S. Beeren, H. Liang, S. Simonsen, C. Schmitz, T. Huber, B.A. Messerle, G. Otting, A Dipicolinic Acid Tag for Rigid Lanthanide Tagging of Proteins and Paramagnetic NMR Spectroscopy, *JACS*, 130 (2008) 10486-10487.
- [11] G. Pintacuda, A.Y. Park, M.A. Keniry, N.E. Dixon, G. Otting, Lanthanide Labeling Offers Fast NMR Approach to 3D Structure Determinations of Protein-Protein Complexes, *JACS*, 128 (2006) 3696-3702.
- [12] K.D. Brewer, T. Bacaj, A. Cavalli, C. Camilloni, J.D. Swarbrick, J. Liu, A. Zhou, P. Zhou, N. Barlow, J. Xu, A.B. Seven, E.A. Prinslow, R. Voleti, D. Häussinger, A.M.J.J. Bonvin, D.R. Tomchick, M. Vendruscolo, B. Graham, T.C. Südhof, J. Rizo, Dynamic binding mode of a Synaptotagmin-1-SNARE complex in solution, *Nat. Struct. Mol. Biol.*, 22 (2015) 555.
- [13] V. Gaponenko, A.S. Altieri, J. Li, R.A. Byrd, Breaking symmetry in the structure determination of (large) symmetric protein dimers, *J. Biomol. NMR*, 24 (2002) 143-148.
- [14] F. Peters, M. Maestre-Martinez, A. Leonov, L. Kovacic, S. Becker, R. Boelens, C. Griesinger, Cys-Ph-TAHA: a lanthanide binding tag for RDC and PCS enhanced protein NMR, *J. Biomol. NMR*, 51 (2011) 329-337.
- [15] J. Wöhnert, K.J. Franz, M. Nitz, B. Imperiali, H. Schwalbe, Protein Alignment by a Coexpressed Lanthanide-Binding Tag for the Measurement of Residual Dipolar Couplings, *JACS*, 125 (2003) 13338-13339.
- [16] I. Bertini, P. Turano, A.J. Vila, Nuclear magnetic resonance of paramagnetic metalloproteins, *Chem. Rev.*, 93 (1993) 2833-2932.
- [17] J. Iwahara, G.M. Clore, Detecting transient intermediates in macromolecular binding by paramagnetic NMR, *Nature*, 440 (2006) 1227.
- [18] I. Bertini, C. Luchinat, G. Parigi, R. Pierattelli, NMR Spectroscopy of Paramagnetic Metalloproteins, *ChemBioChem*, 6 (2005) 1536-1549.
- [19] L. Banci, I. Bertini, K.L. Bren, M.A. Cremonini, H.B. Gray, C. Luchinat, P. Turano, The use of pseudocontact shifts to refine solution structures of paramagnetic metalloproteins: Met80Ala cyano-cytochrome c as an example, *J. Biol. Inorg. Chem.*, 1 (1996) 117-126.
- [20] G. Pintacuda, M. John, X.-C. Su, G. Otting, NMR Structure Determination of Protein-Ligand Complexes by Lanthanide Labeling, *Acc. Chem. Res.*, 40 (2007) 206-212.
- [21] T. Ikegami, L. Verdier, P. Sakhaei, S. Grimme, B. Pescatore, K. Saxena, K.M. Fiebig, C. Griesinger, Novel Techniques for Weak Alignment of Proteins in Solution Using Chemical Tags Coordinating Lanthanide Ions, *J. Biomol. NMR*, 29 (2004) 339-349.
- [22] G.M. Clore, J. Iwahara, Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes, *Chem. Rev.*, 109 (2009) 4108-4139.
- [23] D. Joss, R.M. Walliser, K. Zimmermann, D. Häussinger, Conformationally locked lanthanide chelating tags for convenient pseudocontact shift protein nuclear magnetic resonance spectroscopy, *J. Biomol. NMR*, (2018) 29-38.

- [24] P.H. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand, M. Ubbink, Design, synthesis, and evaluation of a lanthanide chelating protein probe: CLaNP-5 yields predictable paramagnetic effects independent of environment, *JACS*, 130 (2008) 14802-14812.
- [25] M.D. Lee, C.T. Loh, J. Shin, S. Chhabra, M.L. Dennis, G. Otting, J.D. Swarbrick, B. Graham, Compact, hydrophilic, lanthanide-binding tags for paramagnetic NMR spectroscopy, *Chem. Sci.*, 6 (2015) 2614-2624.
- [26] B.J.G. Pearce, S. Jabar, C.-T. Loh, M. Szabo, B. Graham, G. Otting, Structure restraints from heteronuclear pseudocontact shifts generated by lanthanide tags at two different sites, *J. Biomol. NMR*, 68 (2017) 19-32.
- [27] M. Regueiro-Figueroa, B. Bensenane, E. Ruscsák, D. Esteban-Gómez, L.J. Charbonnière, G. Tircsó, I. Tóth, A.d. Blas, T. Rodríguez-Blas, C. Platas-Iglesias, Lanthanide dota-like Complexes Containing a Picolinate Pendant: Structural Entry for the Design of LnIII-Based Luminescent Probes, *Inorg. Chem.*, 50 (2011) 4125-4141.
- [28] M. Strickland, C.D. Schwieters, C. Göbl, A.C.L. Opina, M.-P. Strub, R.E. Swenson, O. Vasalatiy, N. Tjandra, Characterizing the magnetic susceptibility tensor of lanthanide-containing polymethylated-DOTA complexes, *J. Biomol. NMR*, 66 (2016) 125-139.
- [29] K. Tu, M. Gochin, Structure Determination by Restrained Molecular Dynamics Using NMR Pseudocontact Shifts as Experimentally Determined Constraints, *JACS*, 121 (1999) 9276-9285.
- [30] Z. Wu, M.D. Lee, T.J. Carruthers, M. Szabo, M.L. Dennis, J.D. Swarbrick, B. Graham, G. Otting, New Lanthanide Tag for the Generation of Pseudocontact Shifts in DNA by Site-Specific Ligation to a Phosphorothioate Group, *Bioconjugate Chem.*, 28 (2017) 1741-1748.
- [31] H. Yagi, K.V. Loscha, X.-C. Su, M. Stanton-Cook, T. Huber, G. Otting, Tunable paramagnetic relaxation enhancements by [Gd(DPA)₃]³⁻ for protein structure analysis, *J. Biomol. NMR*, 47 (2010) 143-153.
- [32] B. Liang, J.H. Bushweller, L.K. Tamm, Site-Directed Parallel Spin-Labeling and Paramagnetic Relaxation Enhancement in Structure Determination of Membrane Proteins by Solution NMR Spectroscopy, *JACS*, 128 (2006) 4389-4397.
- [33] J.R. Gillespie, D. Shortle, Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. distance restraints from paramagnetic relaxation and calculation of an ensemble of structures¹Edited by P. E. Wright, *J. Mol. Biol.*, 268 (1997) 170-184.
- [34] J. Iwahara, C.D. Schwieters, G.M. Clore, Ensemble Approach for NMR Structure Refinement against ¹H Paramagnetic Relaxation Enhancement Data Arising from a Flexible Paramagnetic Group Attached to a Macromolecule, *JACS*, 126 (2004) 5879-5896.
- [35] J. Iwahara, C. Tang, G. Marius Clore, Practical aspects of ¹H transverse paramagnetic relaxation enhancement measurements on macromolecules, *J. Magn. Reson.*, 184 (2007) 185-195.
- [36] L.W. Donaldson, N.R. Skrynnikov, W.-Y. Choy, D.R. Muhandiram, B. Sarkar, J.D. Forman-Kay, L.E. Kay, Structural Characterization of Proteins with an Attached ATCUN Motif by Paramagnetic Relaxation Enhancement NMR Spectroscopy, *JACS*, 123 (2001) 9843-9847.
- [37] G.M. Clore, M.R. Starich, A.M. Gronenborn, Measurement of Residual Dipolar Couplings of Macromolecules Aligned in the Nematic Phase of a Colloidal Suspension of Rod-Shaped Viruses, *JACS*, 120 (1998) 10571-10572.
- [38] M.W.F. Fischer, J.A. Losonczi, J.L. Weaver, J.H. Prestegard, Domain Orientation and Dynamics in Multidomain Proteins from Residual Dipolar Couplings, *Biochemistry*, 38 (1999) 9013-9022.
- [39] J.R. Tolman, H.M. Al-Hashimi, L.E. Kay, J.H. Prestegard, Structural and Dynamic Analysis of Residual Dipolar Coupling Data for Proteins, *JACS*, 123 (2001) 1416-1424.
- [40] J. Meiler, J.J. Prompers, W. Peti, C. Griesinger, R. Brüschweiler, Model-Free Approach to the Dynamic Interpretation of Residual Dipolar Couplings in Globular Proteins, *JACS*, 123 (2001) 6098-6107.
- [41] M. Blackledge, Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings, *Prog. Nucl. Magn. Reson. Spectrosc.*, 46 (2005) 23-61.
- [42] P. Dosset, J.-C. Hus, D. Marion, M. Blackledge, A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings, *J. Biomol. NMR*, 20 (2001) 223-231.
- [43] W. Peti, J. Meiler, R. Brüschweiler, C. Griesinger, Model-Free Analysis of Protein Backbone Motion from Residual Dipolar Couplings, *JACS*, 124 (2002) 5822-5833.
- [44] J.R. Tolman, K. Ruan, NMR Residual Dipolar Couplings as Probes of Biomolecular Dynamics, *Chem. Rev.*, 106 (2006) 1720-1736.
- [45] P. Bernadó, C.W. Bertoncini, C. Griesinger, M. Zweckstetter, M. Blackledge, Defining Long-Range Order and Local Disorder in Native α -Synuclein Using Residual Dipolar Couplings, *JACS*, 127 (2005) 17968-17969.
- [46] K. Chen, N. Tjandra, The Use of Residual Dipolar Coupling in Studying Proteins by NMR, *Top. Curr. Chem.*, 326 (2012) 47-67.

- [47] R. Barbieri, I. Bertini, G. Cavallaro, Y.-M. Lee, C. Luchinat, A. Rosato, Paramagnetically Induced Residual Dipolar Couplings for Solution Structure Determination of Lanthanide Binding Proteins, *JACS*, 124 (2002) 5581-5587.
- [48] T. Saio, F. Inagaki, Structural Study of Proteins by Paramagnetic Lanthanide Probe Methods, in: T.N.M.R.S.o. Japan (Ed.) *Experimental Approaches of NMR Spectroscopy: Methodology and Application to Life Science and Materials Science*, Springer Singapore, Singapore, 2018, pp. 227-252.
- [49] C.D. Barry, A.C.T. North, J.A. Glasel, R.J.P. Williams, A.V. Xavier, Quantitative Determination of Mononucleotide Conformations in Solution using Lanthanide Ion Shift and Broadening NMR Probes, *Nature*, 232 (1971) 236-245.
- [50] B. Bleaney, C.M. Dobson, B.A. Levine, R.B. Martin, R.J.P. Williams, A.V. Xavier, Origin of lanthanide nuclear magnetic resonance shifts and their uses, *J. Chem. Soc., Chem. Commun.*, (1972) 791b-793.
- [51] B. Bleaney, Nuclear magnetic resonance shifts in solution due to lanthanide ions, *J. Magn. Reson.*, 8 (1972) 91-100.
- [52] A.M. Funk, K.-L.N.A. Finney, P. Harvey, A.M. Kenwright, E.R. Neil, N.J. Rogers, P. Kanthi Senanayake, D. Parker, Critical analysis of the limitations of Bleaney's theory of magnetic anisotropy in paramagnetic lanthanide coordination complexes, *Chem. Sci.*, 6 (2015) 1655-1662.
- [53] E.A. Suturina, K. Mason, C.F.G.C. Geraldes, I. Kuprov, D. Parker, Beyond Bleaney's Theory: Experimental and Theoretical Analysis of Periodic Trends in Lanthanide-Induced Chemical Shift, *Angew. Chem. Int. Ed.*, 56 (2017) 12215-12218.
- [54] C.C. Hinckley, Paramagnetic shifts in solutions of cholesterol and the dipyrindine adduct of trisdipivalomethanatoeuropium(III). A shift reagent, *JACS*, 91 (1969) 5160-5162.
- [55] P.V. Demarco, T.K. Elzey, R.B. Lewis, E. Wenkert, Tris(dipivalomethanato)europium(III). Shift reagent for use in the proton magnetic resonance analysis of steroids and terpenoids, *JACS*, 92 (1970) 5737-5739.
- [56] A.F. Cockerill, D.M. Rackham, Quantitation of the chemical shifts induced by tris(dipivalomethanato)europium III in the PMR spectra of hydroxyadamantanes and cyclopentanol, *Tetrahedron Lett.*, 11 (1970) 5149-5152.
- [57] G.M. Whitesides, D.W. Lewis, Tris[3-(tert-butylhydroxymethylene)-d-camphorato]europium(III). A reagent for determining enantiomeric purity, *JACS*, 92 (1970) 6979-6980.
- [58] H.L. Goering, J.N. Eikenberry, G.S. Koerner, Tris[3-(trifluoromethylhydroxymethylene)-d-camphorato]europium(III). Chiral shift reagent for direct determination of enantiomeric compositions, *JACS*, 93 (1971) 5913-5914.
- [59] T. Saio, K. Ogura, M. Yokochi, Y. Kobashigawa, F. Inagaki, Two-point anchoring of a lanthanide-binding peptide to a target protein enhances the paramagnetic anisotropic effect, *J. Biomol. NMR*, 44 (2009) 157-166.
- [60] K.J. Franz, M. Nitz, B. Imperiali, Lanthanide-Binding Tags as Versatile Protein Coexpression Probes, *ChemBioChem*, 4 (2003) 265-271.
- [61] M. Nitz, K.J. Franz, R.L. Maglathlin, B. Imperiali, A Powerful Combinatorial Screen to Identify High-Affinity Terbium(III)-Binding Peptides, *ChemBioChem*, 4 (2003) 272-276.
- [62] A. Leonov, B. Voigt, F. Rodriguez-Castañeda, P. Sakhaii, C. Griesinger, Convenient Synthesis of Multifunctional EDTA-Based Chiral Metal Chelates Substituted with an S-Mesylcysteine, *Chem. Eur. J.*, 11 (2005) 3342-3348.
- [63] B. Graham, C.T. Loh, J.D. Swarbrick, P. Ung, J. Shin, H. Yagi, X. Jia, S. Chhabra, N. Barlow, G. Pintacuda, T. Huber, G. Otting, DOTA-Amide Lanthanide Tag for Reliable Generation of Pseudocontact Shifts in Protein NMR Spectra, *Bioconjugate Chem.*, 22 (2011) 2118-2125.
- [64] M. Gochin, H. Roder, Protein structure refinement based on paramagnetic NMR shifts: applications to wild-type and mutant forms of cytochrome c, *Protein Sci.*, 4 (1995) 296-305.
- [65] Z. Gong, C.D. Schwieters, C. Tang, Theory and practice of using solvent paramagnetic relaxation enhancement to characterize protein conformational dynamics, *Methods*, (2018).
- [66] J. Gao, E. Liang, R. Ma, F. Li, Y. Liu, J. Liu, L. Jiang, C. Li, H. Dai, J. Wu, X. Su, W. He, K. Ruan, Fluorine Pseudocontact Shifts Used for Characterizing the Protein-Ligand Interaction Mode in the Limit of NMR Intermediate Exchange, *Angew. Chem. Int. Ed.*, 56 (2017) 12982-12986.
- [67] U. Brath, S.I. Swamy, A.X. Veiga, C.-C. Tung, F. Van Petegem, M. Erdélyi, Paramagnetic Ligand Tagging To Identify Protein Binding Sites, *JACS*, 137 (2015) 11391-11398.
- [68] M. John, G. Pintacuda, A.Y. Park, N.E. Dixon, G. Otting, Structure Determination of Protein-Ligand Complexes by Transferred Paramagnetic Shifts, *JACS*, 128 (2006) 12910-12916.
- [69] J.-Y. Guan, P.H.J. Keizers, W.-M. Liu, F. Löhner, S.P. Skinner, E.A. Heeneman, H. Schwalbe, M. Ubbink, G. Siegal, Small-Molecule Binding Sites on Proteins Established by Paramagnetic NMR Spectroscopy, *JACS*, 135 (2013) 5859-5868.

- [70] T. Saio, K. Ogura, K. Shimizu, M. Yokochi, T.R. Burke, F. Inagaki, An NMR strategy for fragment-based ligand screening utilizing a paramagnetic lanthanide probe, *J. Biomol. NMR*, 51 (2011) 395.
- [71] K. Zimmermann, D. Joss, T. Müntener, E.S. Nogueira, M. Schäfer, L. Knörr, F.W. Monnard, D. Häussinger, Localization of ligands within human carbonic anhydrase II using ¹⁹F pseudocontact shift analysis, *Chem. Sci.*, 10 (2019) 5064-5072.
- [72] C. Öster, S. Kosol, C. Hartlmüller, J.M. Lamley, D. Iuga, A. Oss, M.-L. Org, K. Vanatalu, A. Samoson, T. Madl, J.R. Lewandowski, Characterization of Protein–Protein Interfaces in Large Complexes by Solid-State NMR Solvent Paramagnetic Relaxation Enhancements, *JACS*, 139 (2017) 12165-12174.
- [73] C. Göbl, M. Resch, M. Strickland, C. Hartlmüller, M. Viertler, N. Tjandra, T. Madl, Increasing the Chemical-Shift Dispersion of Unstructured Proteins with a Covalent Lanthanide Shift Reagent, *Angew. Chem. Int. Ed.*, 55 (2016) 14847-14851.
- [74] A. Canales, I. Boos, L. Perkams, L. Karst, T. Luber, T. Karagiannis, G. Domínguez, F.J. Cañada, J. Pérez-Castells, D. Häussinger, C. Unverzagt, J. Jiménez-Barbero, Breaking the Limits in Analyzing Carbohydrate Recognition by NMR Spectroscopy: Resolving Branch-Selective Interaction of a Tetra-Antennary N-Glycan with Lectins, *Angew. Chem. Int. Ed.*, 56 (2017) 14987-14991.
- [75] M. Vonci, K. Mason, E.A. Suturina, A.T. Frawley, S.G. Worswick, I. Kuprov, D. Parker, E.J.L. McInnes, N.F. Chilton, Rationalization of Anomalous Pseudocontact Shifts and Their Solvent Dependence in a Series of C₃-Symmetric Lanthanide Complexes, *JACS*, 139 (2017) 14166-14172.
- [76] A. Roca-Sabio, M. Regueiro-Figueroa, D. Esteban-Gómez, A. de Blas, T. Rodríguez-Blas, C. Platas-Iglesias, Density functional dependence of molecular geometries in lanthanide(III) complexes relevant to bioanalytical and biomedical applications, *Comput. Theor. Chem.*, 999 (2012) 93-104.
- [77] G. Hong, M. Dolg, L. Li, A comparison of scalar-relativistic ZORA and DKH density functional schemes: monohydrides, monooxides and monofluorides of La, Lu, Ac and Lr, *Chem. Phys. Lett.*, 334 (2001) 396-402.
- [78] M. Purgel, Z. Baranyai, A. de Blas, T. Rodríguez-Blas, I. Bánya, C. Platas-Iglesias, I. Tóth, An NMR and DFT Investigation on the Conformational Properties of Lanthanide(III) 1,4,7,10-Tetraazacyclododecane-1,4,7,10-tetraacetate Analogues Containing Methylenephosphonate Pendant Arms, *Inorg. Chem.*, 49 (2010) 4370-4382.
- [79] G. Schreckenbach, T. Ziegler, Calculation of the G-Tensor of Electron Paramagnetic Resonance Spectroscopy Using Gauge-Including Atomic Orbitals and Density Functional Theory, *J. Phys. Chem. A*, 101 (1997) 3388-3399.
- [80] S. Per, P.E.A. Turchi, A. Landa, V. Lordi, Ground-state properties of rare-earth metals: an evaluation of density-functional theory, *J. Phys. Condens. Matter*, 26 (2014) 416001.
- [81] X. Dai, Y. Gao, M. Xin, Z. Wang, R. Zhou, The ground state and electronic structure of Gd@C₈₂: a systematic theoretical investigation of first principle density functionals, *J. Chem. Phys.*, 141 (2014) 244306.
- [82] S. Grimm, G. Schoendorff, A.K. Wilson, Gauging the Performance of Density Functionals for Lanthanide-Containing Molecules, *J. Chem. Theory Comput.*, 12 (2016) 1259-1266.
- [83] L.E. Aebbersold, S.H. Yuwono, G. Schoendorff, A.K. Wilson, Efficacy of Density Functionals and Relativistic Effective Core Potentials for Lanthanide-Containing Species: The Ln₅₄ Molecule Set, *J. Chem. Theory Comput.*, 13 (2017) 2831-2839.
- [84] L. Cerofolini, J.M. Silva, E. Ravera, M. Romanelli, C.F.G.C. Geraldes, A.L. Macedo, M. Fragai, G. Parigi, C. Luchinat, How Do Nuclei Couple to the Magnetic Moment of a Paramagnetic Center? A New Theory at the Gauntlet of the Experiments, *J. Phys. Chem. Lett.*, 10 (2019) 3610-3614.
- [85] J. Koehler, J. Meiler, Expanding the utility of NMR restraints with paramagnetic compounds: Background and practical aspects, *Prog. Nucl. Magn. Reson. Spectrosc.*, 59 (2011) 360-389.
- [86] F. Rodríguez-Castañeda, P. Haberz, A. Leonov, C. Griesinger, Paramagnetic tagging of diamagnetic proteins for solution NMR, *Magn. Reson. Chem.*, 44 (2006) S10-S16.
- [87] X.-C. Su, G. Otting, Paramagnetic labelling of proteins and oligonucleotides for NMR, *J. Biomol. NMR*, 46 (2010) 101-112.
- [88] P.H.J. Keizers, M. Ubbink, Paramagnetic tagging for protein structure and dynamics analysis, *Prog. Nucl. Magn. Reson. Spectrosc.*, 58 (2011) 88-96.
- [89] W.-M. Liu, M. Overhand, M. Ubbink, The application of paramagnetic lanthanoid ions in NMR spectroscopy on proteins, *Coord. Chem. Rev.*, 273-274 (2014) 2-12.
- [90] K.N. Allen, B. Imperiali, Lanthanide-tagged proteins—an illuminating partnership, *Curr. Opin. Chem. Biol.*, 14 (2010) 247-254.
- [91] T. Saio, K. Ishimori, Accelerating structural life science by paramagnetic lanthanide probe methods, *Biochim. Biophys. Acta. Gen. Subj.*, (2019).

- [92] X.-C. Su, J.-L. Chen, Site-Specific Tagging of Proteins with Paramagnetic Ions for Determination of Protein Structures in Solution and in Cells, *Acc. Chem. Res.*, 52 (2019) 1675-1686.
- [93] M. John, G. Otting, Strategies for Measurements of Pseudocontact Shifts in Protein NMR Spectroscopy, *Chemphyschem*, 8 (2007) 2309-2313.
- [94] G. Otting, Prospects for lanthanides in structural biology by NMR, *J. Biomol. NMR*, 42 (2008) 1-9.
- [95] M.A.S. Hass, M. Ubbink, Structure determination of protein–protein complexes with long-range anisotropic paramagnetic NMR restraints, *Curr. Opin. Struct. Biol.*, 24 (2014) 45-53.
- [96] C. Göbl, T. Madl, B. Simon, M. Sattler, NMR approaches for structural analysis of multidomain proteins and complexes in solution, *Prog. Nucl. Magn. Reson. Spectrosc.*, 80 (2014) 26-63.
- [97] A. Carlon, E. Ravera, W. Andrałojć, G. Parigi, G.N. Murshudov, C. Luchinat, How to tackle protein structural data from solution and solid state: An integrated approach, *Prog. Nucl. Magn. Reson. Spectrosc.*, 92-93 (2016) 54-70.
- [98] A. Bhaumik, C. Luchinat, G. Parigi, E. Ravera, M. Rinaldelli, NMR crystallography on paramagnetic systems: solved and open issues, *CrystEngComm*, 15 (2013) 8639-8656.
- [99] J.H. Prestegard, D.A. Agard, K.W. Moremen, L.A. Lavery, L.C. Morris, K. Pederson, Sparse labeling of proteins: Structural characterization from long range constraints, *J. Magn. Reson.*, 241 (2014) 32-40.
- [100] K. Kato, T. Yamaguchi, Paramagnetic NMR probes for characterization of the dynamic conformations and interactions of oligosaccharides, *Glycoconj. J.*, 32 (2015) 505-513.
- [101] I. Bertini, C. Luchinat, G. Parigi, Magnetic susceptibility in paramagnetic NMR, *Prog. Nucl. Magn. Reson. Spectrosc.*, 40 (2002) 249-273.
- [102] A.J. Pell, G. Pintacuda, C.P. Grey, Paramagnetic NMR in solution and the solid state, *Prog. Nucl. Magn. Reson. Spectrosc.*, 111 (2019) 1-271.
- [103] C. Schmitz, M.J. Stanton-Cook, X.C. Su, G. Otting, T. Huber, Numbat: an interactive software tool for fitting Deltachi-tensors to molecular coordinates using pseudocontact shifts, *J. Biomol. NMR*, 41 (2008) 179-189.
- [104] L. Yao, B. Vögeli, J. Ying, A. Bax, NMR Determination of Amide N–H Equilibrium Bond Length from Concerted Dipolar Coupling Measurements, *JACS*, 130 (2008) 16518-16520.
- [105] K.B. Pilla, G. Otting, T. Huber, Pseudocontact Shift-Driven Iterative Resampling for 3D Structure Determinations of Large Proteins, *J. Mol. Biol.*, 428 (2016) 522-532.
- [106] G.M. Clore, Accurate and rapid docking of protein–protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization, *PNAS*, 97 (2000) 9021.
- [107] J.H. Prestegard, H.M. Al-Hashimi, J.R. Tolman, NMR structures of biomolecules using field oriented media and residual dipolar couplings, *Q. Rev. Biophys.*, 33 (2000) 371-424.
- [108] C.A. Bewley, G.M. Clore, Determination of the Relative Orientation of the Two Halves of the Domain-Swapped Dimer of Cyanovirin-N in Solution Using Dipolar Couplings and Rigid Body Minimization, *JACS*, 122 (2000) 6009-6016.
- [109] G.-W. Li, H. Liu, F. Qiu, X.-J. Wang, X.-X. Lei, Residual Dipolar Couplings in Structure Determination of Natural Products, *Nat. Prod. Bioprospect.*, 8 (2018) 279-295.
- [110] D. MacDonald, P. Lu, Residual dipolar couplings in nucleic acid structure determination, *Curr. Opin. Struct. Biol.*, 12 (2002) 337-343.
- [111] R.D. Shannon, Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, *Acta Crystallogr. A.*, 32 (1976) 751-767.
- [112] I. Bertini, M.B.L. Janik, Y.-M. Lee, C. Luchinat, A. Rosato, Magnetic Susceptibility Tensor Anisotropies for a Lanthanide Ion Series in a Fixed Protein Matrix, *JACS*, 123 (2001) 4181-4188.
- [113] I. Bertini, A. Donaire, B. Jimenez, C. Luchinat, G. Parigi, M. Piccioli, L. Poggi, Paramagnetism-based versus classical constraints: an analysis of the solution structure of Ca Ln calbindin D9k, *J. Biomol. NMR*, 21 (2001) 85-98.
- [114] C. Ma, S.J. Opella, Lanthanide Ions Bind Specifically to an Added “EF-Hand” and Orient a Membrane Protein in Micelles for Solution NMR Spectroscopy, *J. Magn. Reson.*, 146 (2000) 381-384.
- [115] M. Nitz, M. Sherawat, K.J. Franz, E. Peisach, K.N. Allen, B. Imperiali, Structural Origin of the High Affinity of a Chemically Evolved Lanthanide-Binding Peptide, *Angew. Chem. Int. Ed.*, 43 (2004) 3682-3685.
- [116] T. Saio, M. Yokochi, H. Kumeta, F. Inagaki, PCS-based structure determination of protein–protein complexes, *J. Biomol. NMR*, 46 (2010) 271-280.
- [117] Y. Kobashigawa, T. Saio, M. Ushio, M. Sekiguchi, M. Yokochi, K. Ogura, F. Inagaki, Convenient method for resolving degeneracies due to symmetry of the magnetic susceptibility tensor and its application to pseudo contact shift-based protein–protein complex structure determination, *J. Biomol. NMR*, 53 (2012) 53-63.
- [118] X.-C. Su, K. McAndrew, T. Huber, G. Otting, Lanthanide-Binding Peptides for NMR Measurements of Residual Dipolar Couplings and Paramagnetic Effects from Multiple Angles, *JACS*, 130 (2008) 1681-1687.

- [119] X.-C. Su, T. Huber, N.E. Dixon, G. Otting, Site-Specific Labelling of Proteins with a Rigid Lanthanide-Binding Tag, *ChemBioChem*, 7 (2006) 1599-1604.
- [120] M. Prudêncio, J. Rohovec, J.A. Peters, E. Tocheva, M.J. Boulanger, M.E.P. Murphy, H.-J. Hupkes, W. Kusters, A. Impagliazzo, M. Ubbink, A Caged Lanthanide Complex as a Paramagnetic Shift Agent for Protein NMR, *Chem. Eur. J.*, 10 (2004) 3252-3260.
- [121] D. Parker, R.S. Dickins, H. Puschmann, C. Crossland, J.A.K. Howard, Being Excited by Lanthanide Coordination Complexes: Aqua Species, Chirality, Excited-State Chemistry, and Exchange Dynamics, *Chem. Rev.*, 102 (2002) 1977-2010.
- [122] G. Pintacuda, A. Moshref, A. Leonchiks, A. Sharipo, G. Otting, Site-specific Labelling with a Metal Chelator for Protein-structure Refinement, *J. Biomol. NMR*, 29 (2004) 351-361.
- [123] A. Dvoretzky, V. Gaponenko, P.R. Rosevear, Derivation of structural restraints using a thiol-reactive chelator, *FEBS Lett.*, 528 (2002) 189-192.
- [124] P. Haberz, F. Rodriguez-Castañeda, J. Junker, S. Becker, A. Leonov, C. Griesinger, Two New Chiral EDTA-Based Metal Chelates for Weak Alignment of Proteins in Solution, *Org. Lett.*, 8 (2006) 1275-1278.
- [125] X.-C. Su, B. Man, S. Beeren, H. Liang, S. Simonsen, C. Schmitz, T. Huber, B.A. Messerle, G. Otting, A Dipicolinic Acid Tag for Rigid Lanthanide Tagging of Proteins and Paramagnetic NMR Spectroscopy, *JACS*, 130 (2008) 10486-10487.
- [126] B. Man, X.-C. Su, H. Liang, S. Simonsen, T. Huber, B.A. Messerle, G. Otting, 3-Mercapto-2,6-Pyridinedicarboxylic Acid: A Small Lanthanide-Binding Tag for Protein Studies by NMR Spectroscopy, *Chem. Eur. J.*, 16 (2010) 3827-3832.
- [127] J.D. Swarbrick, P. Ung, X.-C. Su, A. Maleckis, S. Chhabra, T. Huber, G. Otting, B. Graham, Engineering of a bis-chelator motif into a protein α -helix for rigid lanthanide binding and paramagnetic NMR spectroscopy, *Chem. Commun.*, 47 (2011) 7368-7370.
- [128] J.D. Swarbrick, P. Ung, S. Chhabra, B. Graham, An Iminodiacetic Acid Based Lanthanide Binding Tag for Paramagnetic Exchange NMR Spectroscopy, *Angew. Chem. Int. Ed.*, 50 (2011) 4403-4406.
- [129] H. Yagi, A. Maleckis, G. Otting, A systematic study of labelling an α -helix in a protein with a lanthanide using IDA-SH or NTA-SH tags, *J. Biomol. NMR*, 55 (2013) 157-166.
- [130] X. Jia, A. Maleckis, T. Huber, G. Otting, 4,4'-Dithiobisdipicolinic Acid: A Small and Convenient Lanthanide Binding Tag for Protein NMR Spectroscopy, *Chem. Eur. J.*, 17 (2011) 6830-6836.
- [131] Q.-F. Li, Y. Yang, A. Maleckis, G. Otting, X.-C. Su, Thiol-ene reaction: a versatile tool in site-specific labelling of proteins with chemically inert tags for paramagnetic NMR, *Chem. Commun.*, 48 (2012) 2704-2706.
- [132] Y. Yang, Q.-F. Li, C. Cao, F. Huang, X.-C. Su, Site-Specific Labeling of Proteins with a Chemically Stable, High-Affinity Tag for Protein Study, *Chem. Eur. J.*, 19 (2013) 1097-1103.
- [133] F. Huang, Y.-Y. Pei, H.-H. Zuo, J.-L. Chen, Y. Yang, X.-C. Su, Bioconjugation of Proteins with a Paramagnetic NMR and Fluorescent Tag, *Chem. Eur. J.*, 19 (2013) 17141-17149.
- [134] Y. Yang, J.-T. Wang, Y.-Y. Pei, X.-C. Su, Site-specific tagging proteins via a rigid, stable and short thioether tether for paramagnetic spectroscopic analysis, *Chem. Commun.*, 51 (2015) 2824-2827.
- [135] C.-T. Loh, B. Graham, E.H. Abdelkader, K.L. Tuck, G. Otting, Generation of Pseudocontact Shifts in Proteins with Lanthanides Using Small "Clickable" Nitrilotriacetic Acid and Iminodiacetic Acid Tags, *Chem. Eur. J.*, 21 (2015) 5084-5092.
- [136] W.-X. Jiang, X.-H. Gu, X. Dong, C. Tang, Lanthanoid tagging via an unnatural amino acid for protein structure characterization, *J. Biomol. NMR*, 67 (2017) 273-282.
- [137] J.-L. Chen, Y. Zhao, Y.-J. Gong, B.-B. Pan, X. Wang, X.-C. Su, Stable and rigid DTPA-like paramagnetic tags suitable for in vitro and in situ protein NMR analysis, *J. Biomol. NMR*, 70 (2018) 77-92.
- [138] M.F. Tweedle, J.J. Hagan, K. Kumar, S. Mantha, C.A. Chang, Reaction of gadolinium chelates with endogenously available ions, *Magn. Reson. Imaging*, 9 (1991) 409-415.
- [139] G.J. Stasiuk, N.J. Long, The ubiquitous DOTA and its derivatives: the impact of 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid on biomedical imaging, *Chem. Commun.*, 49 (2013) 2732-2746.
- [140] H.R. Maecke, J.C. Reubi, Somatostatin Receptors as Targets for Nuclear Medicine Imaging and Radionuclide Treatment, *J. Nucl. Med.*, 52 (2011) 841-844.
- [141] D. Häussinger, S. Grzesiek, A new lanthanide-chelating tag with very high metal affinity was developed for protein NMR on a calcium binding protein, *Conference Abstract Book*, 1 (2005) 1-10.
- [142] A.C.L. Opina, M. Strickland, Y.-S. Lee, N. Tjandra, R.A. Byrd, R.E. Swenson, O. Vasalatiy, Analysis of the isomer ratios of polymethylated-DOTA complexes and the implications on protein structural studies, *Dalton Trans.*, 45 (2016) 4673-4687.

- [143] E.A. Suturina, D. Haussinger, K. Zimmermann, L. Garbuio, M. Yulikov, G. Jeschke, I. Kuprov, Model-free extraction of spin label position distributions from pseudocontact shift data, *Chem. Sci.*, 8 (2017) 2751-2757.
- [144] V.S. Mironov, Y.G. Galyametdinov, A. Ceulemans, C. Görlner-Walrand, K. Binnemans, Influence of crystal-field perturbations on the room-temperature magnetic anisotropy of lanthanide complexes, *Chem. Phys. Lett.*, 345 (2001) 132-140.
- [145] M.D. Lee, M.L. Dennis, B. Graham, J.D. Swarbrick, Short two-armed lanthanide-binding tags for paramagnetic NMR spectroscopy based on chiral 1,4,7,10-tetrakis(2-hydroxypropyl)-1,4,7,10-tetraazacyclododecane scaffolds, *Chem. Commun.*, 53 (2017) 13205-13208.
- [146] M.D. Lee, M.L. Dennis, J.D. Swarbrick, B. Graham, Enantiomeric two-armed lanthanide-binding tags for complementary effects in paramagnetic NMR spectroscopy, *Chem. Commun.*, 52 (2016) 7954-7957.
- [147] T. Müntener, J. Kottelat, A. Huber, D. Häussinger, New Lanthanide Chelating Tags for PCS NMR Spectroscopy with Reduction Stable, Rigid Linkers for Fast and Irreversible Conjugation to Proteins, *Bioconjugate Chem.*, 29 (2018) 3344-3351.
- [148] D. Joss, M.-S. Bertrams, D. Häussinger, A sterically overcrowded, isopropyl-substituted lanthanide chelating tag for protein PCS NMR spectroscopy: Synthesis of its macrocyclic scaffold and benchmarking on ubiquitin S57C and hCA II S166C, *Chem. Eur. J.*, in press (2019) doi.org/10.1002/chem.201901692.
- [149] E.A. Suturina, I. Kuprov, Pseudocontact shifts from mobile spin labels, *Phys. Chem. Chem. Phys.*, 18 (2016) 26412-26422.
- [150] D. Joss, D. Häussinger, P4T – A lanthanide chelating tag combining a highly sterically overcrowded backbone with a reductively stable linker, submitted (2019).
- [151] F. Yang, X. Wang, B.-B. Pan, X.-C. Su, Single-armed phenylsulfonated pyridine derivative of DOTA is rigid and stable paramagnetic tag in protein analysis, *Chem. Commun.*, 52 (2016) 11535-11538.
- [152] P.H.J. Keizers, J.F. Desreux, M. Overhand, M. Ubbink, Increased Paramagnetic Effect of a Lanthanide Protein Probe by Two-Point Attachment, *JACS*, 129 (2007) 9292-9293.
- [153] M.D. Vlasie, C. Comuzzi, A.M.C.H. van den Nieuwendijk, M. Prudêncio, M. Overhand, M. Ubbink, Long-Range-Distance NMR Effects in a Protein Labeled with a Lanthanide-DOTA Chelate, *Chem. Eur. J.*, 13 (2007) 1715-1723.
- [154] W.-M. Liu, S.P. Skinner, M. Timmer, A. Blok, M.A.S. Hass, D.V. Filippov, M. Overhand, M. Ubbink, A Two-Armed Lanthanoid-Chelating Paramagnetic NMR Probe Linked to Proteins via Thioether Linkages, *Chem. Eur. J.*, 20 (2014) 6256-6258.
- [155] W.-M. Liu, P.H.J. Keizers, M.A.S. Hass, A. Blok, M. Timmer, A.J.C. Sarris, M. Overhand, M. Ubbink, A pH-Sensitive, Colorful, Lanthanide-Chelating Paramagnetic NMR Probe, *JACS*, 134 (2012) 17306-17313.
- [156] D. Shishmarev, G. Otting, How reliable are pseudocontact shifts induced in proteins and ligands by mobile paramagnetic metal tags? A modelling study, *J. Biomol. NMR*, 56 (2013) 203-216.
- [157] M.A.S. Hass, P.H.J. Keizers, A. Blok, Y. Hiruma, M. Ubbink, Validation of a Lanthanide Tag for the Analysis of Protein Dynamics by Paramagnetic NMR Spectroscopy, *JACS*, 132 (2010) 9952-9953.
- [158] K. Mason, N.J. Rogers, E.A. Suturina, I. Kuprov, J.A. Aguilar, A.S. Batsanov, D.S. Yufit, D. Parker, PARASHIFT Probes: Solution NMR and X-ray Structural Studies of Macrocyclic Ytterbium and Yttrium Complexes, *Inorg. Chem.*, 56 (2017) 4028-4038.
- [159] A.C. Harnden, E.A. Suturina, A.S. Batsanov, P.K. Senanayake, M.A. Fox, K. Mason, M. Vonci, E.J.L. McInnes, N.F. Chilton, D. Parker, Unravelling the Complexities of Pseudocontact Shift Analysis in Lanthanide Coordination Complexes of Differing Symmetry, *Angew. Chem. Int. Ed.*, 0 (2019).
- [160] P.H.J. Keizers, B. Mersinli, W. Reinle, J. Donauer, Y. Hiruma, F. Hannemann, M. Overhand, R. Bernhardt, M. Ubbink, A Solution Model of the Complex Formed by Adrenodoxin and Adrenodoxin Reductase Determined by Paramagnetic NMR Spectroscopy, *Biochemistry*, 49 (2010) 6846-6855.
- [161] L. de la Cruz, T.H.D. Nguyen, K. Ozawa, J. Shin, B. Graham, T. Huber, G. Otting, Binding of Low Molecular Weight Inhibitors Promotes Large Conformational Changes in the Dengue Virus NS2B-NS3 Protease: Fold Analysis by Pseudocontact Shifts, *JACS*, 133 (2011) 19205-19215.
- [162] I. Bertini, V. Calderone, L. Cerofolini, M. Fragai, C.F.G.C. Geraldès, P. Hermann, C. Luchinat, G. Parigi, J.M.C. Teixeira, The catalytic domain of MMP-1 studied through tagged lanthanides, *FEBS Lett.*, 586 (2012) 557-567.
- [163] J. Li, K.B. Pilla, Q. Li, Z. Zhang, X. Su, T. Huber, J. Yang, Magic Angle Spinning NMR Structure Determination of Proteins from Pseudocontact Shifts, *JACS*, 135 (2013) 8294-8303.
- [164] H. Yagi, K.B. Pilla, A. Maleckis, B. Graham, T. Huber, G. Otting, Three-Dimensional Protein Fold Determination from Backbone Amide Pseudocontact Shifts Generated by Lanthanide Tags at Multiple Sites, *Structure*, 21 (2013) 883-890.

- [165] Y. Hiruma, M.A.S. Hass, Y. Kikui, W.-M. Liu, B. Ölmez, S.P. Skinner, A. Blok, A. Kloosterman, H. Koteishi, F. Löhr, H. Schwalbe, M. Nojiri, M. Ubbink, The Structure of the Cytochrome P450cam–Putidaredoxin Complex Determined by Paramagnetic NMR Spectroscopy and Crystallography, *J. Mol. Biol.*, 425 (2013) 4353–4365.
- [166] W.-N. Chen, K.V. Loscha, C. Nitsche, B. Graham, G. Otting, The dengue virus NS2B–NS3 protease retains the closed conformation in the complex with BPTI, *FEBS Lett.*, 588 (2014) 2206–2211.
- [167] D. Shishmarev, Y. Wang, C.E. Mason, X.-C. Su, A.J. Oakley, B. Graham, T. Huber, N.E. Dixon, G. Otting, Intramolecular binding mode of the C-terminus of Escherichia coli single-stranded DNA binding protein determined by nuclear magnetic resonance spectroscopy, *Nucleic Acids Res.*, 42 (2014) 2750–2757.
- [168] L. de la Cruz, W.-N. Chen, B. Graham, G. Otting, Binding mode of the activity-modulating C-terminal segment of NS2B to NS3 in the dengue virus NS2B–NS3 protease, *FEBS J.*, 281 (2014) 1517–1533.
- [169] A.R. Camacho-Zarco, F. Munari, M. Wegstroth, W.-M. Liu, M. Ubbink, S. Becker, M. Zweckstetter, Multiple Paramagnetic Effects through a Tagged Reporter Protein, *Angew. Chem. Int. Ed.*, 54 (2015) 336–339.
- [170] D.J. Crick, J.X. Wang, B. Graham, J.D. Swarbrick, H.R. Mott, D. Nietlispach, Integral membrane protein structure determination using pseudocontact shifts, *J. Biomol. NMR*, 61 (2015) 197–207.
- [171] Y.-Z. Pan, B. Quade, K.D. Brewer, M. Szabo, J.D. Swarbrick, B. Graham, J. Rizo, Sequence-specific assignment of methyl groups from the neuronal SNARE complex using lanthanide-induced pseudocontact shifts, *J. Biomol. NMR*, 66 (2016) 281–293.
- [172] T. Saio, K. Ogura, H. Kumeta, Y. Kobashigawa, K. Shimizu, M. Yokochi, K. Kodama, H. Yamaguchi, H. Tsujishita, F. Inagaki, Ligand-driven conformational changes of MurD visualized by paramagnetic NMR, *Sci. Rep.*, 5 (2015) 16685.
- [173] J.-L. Chen, Y. Yang, L.-L. Zhang, H. Liang, T. Huber, X.-C. Su, G. Otting, Analysis of the solution conformations of T4 lysozyme by paramagnetic NMR spectroscopy, *Phys. Chem. Chem. Phys.*, 18 (2016) 5850–5859.
- [174] R.S. Ma, Q.F. Li, A.D. Wang, J.H. Zhang, Z.J. Liu, J.H. Wu, X.C. Su, K. Ruan, Determination of pseudocontact shifts of low-populated excited states by NMR chemical exchange saturation transfer, *Phys. Chem. Chem. Phys.*, 18 (2016) 13794–13798.
- [175] M. Lescanne, S.P. Skinner, A. Blok, M. Timmer, L. Cerofolini, M. Fragai, C. Luchinat, M. Ubbink, Methyl group assignment using pseudocontact shifts with PARAssign, *J. Biomol. NMR*, 69 (2017) 183–195.
- [176] M. Lescanne, P. Ahuja, A. Blok, M. Timmer, T. Akerud, M. Ubbink, Methyl group reorientation under ligand binding probed by pseudocontact shifts, *J. Biomol. NMR*, 71 (2018) 275–285.
- [177] D.E. Kamen, S.M. Cahill, M.E. Girvin, Multiple Alignment of Membrane Proteins for Measuring Residual Dipolar Couplings Using Lanthanide Ions Bound to a Small Metal Chelator, *JACS*, 129 (2007) 1846–1847.
- [178] X. Xu, P.H.J. Keizers, W. Reinle, F. Hannemann, R. Bernhardt, M. Ubbink, Intermolecular dynamics studied by paramagnetic tagging, *J. Biomol. NMR*, 43 (2009) 247–254.
- [179] S.C. Chiliveri, J.M. Louis, R. Ghirlando, J.L. Baber, A. Bax, Tilted, Uninterrupted, Monomeric HIV-1 gp41 Transmembrane Helix from Residual Dipolar Couplings, *JACS*, 140 (2018) 34–37.
- [180] C.A. Barnes, Y. Shen, J. Ying, Y. Takagi, D.A. Torchia, J.R. Sellers, A. Bax, Remarkable Rigidity of the Single α -Helical Domain of Myosin-VI As Revealed by NMR Spectroscopy, *JACS*, 141 (2019) 9004–9017.
- [181] W.-N. Chen, C. Nitsche, K.B. Pilla, B. Graham, T. Huber, C.D. Klein, G. Otting, Sensitive NMR Approach for Determining the Binding Mode of Tightly Binding Ligand Molecules to Protein Targets, *JACS*, 138 (2016) 4539–4546.
- [182] D. Xu, B. Li, J. Gao, Z. Liu, X. Niu, G. Nshogoza, J. Zhang, J. Wu, X.-C. Su, W. He, R. Ma, D. Yang, K. Ruan, Ligand Proton Pseudocontact Shifts Determined from Paramagnetic Relaxation Dispersion in the Limit of NMR Intermediate Exchange, *J. Phys. Chem. Lett.*, 9 (2018) 3361–3367.
- [183] A. Mallagaray, A. Canales, G. Domínguez, J. Jiménez-Barbero, J. Pérez-Castells, A rigid lanthanide binding tag for NMR structural analysis of carbohydrates, *Chem. Commun.*, 47 (2011) 7179–7181.
- [184] M. Erdélyi, E. d'Auvergne, A. Navarro-Vázquez, A. Leonov, C. Griesinger, Dynamics of the Glycosidic Bond: Conformational Space of Lactose, *Chem. Eur. J.*, 17 (2011) 9368–9376.
- [185] S. Yamamoto, T. Yamaguchi, M. Erdélyi, C. Griesinger, K. Kato, Paramagnetic Lanthanide Tagging for NMR Conformational Analyses of N-Linked Oligosaccharides, *Chem. Eur. J.*, 17 (2011) 9280–9282.
- [186] Y. Zhang, S. Yamamoto, T. Yamaguchi, K. Kato, Application of Paramagnetic NMR-Validated Molecular Dynamics Simulation to the Analysis of a Conformational Ensemble of a Branched Oligosaccharide, *Molecules*, 17 (2012) 6658–6671.
- [187] Gabriel A. Rabinovich, Diego O. Croci, Regulatory Circuits Mediated by Lectin-Glycan Interactions in Autoimmunity and Cancer, *Immunity*, 36 (2012) 322–335.
- [188] D.N. Hebert, L. Lamriben, E.T. Powers, J.W. Kelly, The intrinsic and extrinsic effects of N-linked glycans on glycoproteostasis, *Nat. Chem. Biol.*, 10 (2014) 902.

- [189] A. Canales, A. Mallagaray, J. Pérez-Castells, I. Boos, C. Unverzagt, S. André, H.-J. Gabius, F.J. Cañada, J. Jiménez-Barbero, Breaking Pseudo-Symmetry in Multiantennary Complex N-Glycans Using Lanthanide-Binding Tags and NMR Pseudo-Contact Shifts, *Angew. Chem. Int. Ed.*, 52 (2013) 13789-13793.
- [190] Á. Canales, Á. Mallagaray, M.Á. Berbís, A. Navarro-Vázquez, G. Domínguez, F.J. Cañada, S. André, H.-J. Gabius, J. Pérez-Castells, J. Jiménez-Barbero, Lanthanide-Chelating Carbohydrate Conjugates Are Useful Tools To Characterize Carbohydrate Conformation in Solution and Sensitive Sensors to Detect Carbohydrate-Protein Interactions, *JACS*, 136 (2014) 8011-8017.
- [191] A. Mallagaray, G. Domínguez, T. Peters, J. Pérez-Castells, A rigid lanthanide binding tag to aid NMR studies of a 70 kDa homodimeric coat protein of human norovirus, *Chem. Commun.*, 52 (2016) 601-604.
- [192] B. Alberts, Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P., *Molecular Biology of the Cell*, 4th edition, Garland Science, New York.
- [193] S.S. Wijmenga, B.N.M. van Buuren, The use of NMR methods for conformational studies of nucleic acids, *Prog. Nucl. Magn. Reson. Spectrosc.*, 32 (1998) 287-387.
- [194] W. Kurt, NMR Studies of Structure and Function of Biological Macromolecules (Nobel Lecture), *Angew. Chem. Int. Ed.*, 42 (2003) 3340-3363.
- [195] M. Gochin, Nuclear Magnetic Resonance Studies of a Paramagnetic Metallo DNA Complex, *JACS*, 119 (1997) 3377-3378.
- [196] M. Gochin, Nuclear Magnetic Resonance Characterization of a Paramagnetic DNA-drug Complex with High Spin Cobalt; Assignment of the ^1H and ^{31}P NMR Spectra, and Determination of Electronic, Spectroscopic and Molecular Properties, *J. Biomol. NMR*, 12 (1998) 243-257.
- [197] K. Jaudzems, X. Jia, H. Yagi, D. Zhulenkova, B. Graham, G. Otting, E. Liepinsh, Structural Basis for 5'-End-Specific Recognition of Single-Stranded DNA by the R3H Domain from Human Smp-2, *J. Mol. Biol.*, 424 (2012) 42-53.
- [198] M. Strickland, J. Catzaro, R. Rajasekaran, M.-P. Strub, C. O'Hern, G.A. Bermejo, M.F. Summers, J. Marchant, N. Tjandra, Long-Range RNA Structural Information via a Paramagnetically Tagged Reporter Protein, *JACS*, 141 (2019) 1430-1434.
- [199] Y. Song, T.J. Meade, A.V. Astashkin, E.L. Klein, J.H. Enemark, A. Raitsimring, Pulsed dipolar spectroscopy distance measurements in biomacromolecules labeled with Gd(III) markers, *J. Magn. Reson.*, 210 (2011) 59-68.
- [200] D. Goldfarb, Gd $^{3+}$ spin labeling for distance measurements by pulse EPR spectroscopy, *Phys. Chem. Chem. Phys.*, 16 (2014) 9685-9699.